

GenCore version 5.1.5
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 2, 2003, 11:21:20 ; Search time 423.993 Seconds

(without alignments)
6795.341 Million cell updates/sec

Title: US-09-719-554-3_COPY-2502_2600

Perfect score: 99
Sequence: 1 tcacccctgagcgcaagc.....tgtrcaagaagcggaaga 99

Scoring table: IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl:
1: gb_ba:*
2: gb_hg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sy:*
12: gb_sy:*
13: gb_un:*
14: gb_vl:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_ph:*
24: em_pl:*
25: em_ro:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vl:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_mus:*
34: em_htg_pln:*
35: em_htg_rnd:*
36: em_htg_mam:*
37: em_htg_vrt:*
38: em_sy:*
39: em_htgo_hum:*
40: em_htgo_mus:*
41: em_htgo_other:*

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	99	100.0	10499	6	AX007980
2	99	100.0	56093	6	AX329572
3	99	100.0	56093	6	HSAC000064
4	99	100.0	149194	9	AC007566
5	95.8	96.8	40205	9	AF045450
6	95.8	96.8	183951	2	AC027500
7	94.2	95.2	104853	9	AC117444
8	94.2	95.2	142742	9	AF121782
9	94.2	95.2	164988	2	AC080070
10	94.2	95.2	169286	9	AC073066
11	94.2	95.2	169462	2	AC024033
12	94.2	95.2	177720	9	AC019193
13	94.2	95.2	186723	2	AC027752
14	94.2	95.2	340000	9	HS21C080
15	93.2	94.1	120450	9	AC093393
16	92.6	93.5	8339	9	AL162912
17	92.6	93.5	46340	6	AX008043
18	92.6	93.5	88502	2	AC021967
19	92.6	93.5	113189	9	AC007883
20	92.6	93.5	128468	9	AL139090
21	92.6	93.5	135873	2	AC012547
22	92.6	93.5	148724	2	AC010778
23	92.6	93.5	150955	9	AL356111
24	92.6	93.5	156818	9	AC104182
25	92.6	93.5	162887	9	AC097635
26	92.6	93.5	167366	2	AC021774
27	92.6	93.5	169747	9	AL355352
28	92.6	93.5	174228	2	AC016105
29	92.6	93.5	176879	2	AC016709
30	92.6	93.5	177163	9	HS134E15
31	92.6	93.5	187399	9	AC067942
32	92.6	93.5	192815	9	AC013444
33	92.6	93.5	199520	2	AC110420
34	92.6	93.5	219476	9	AC068945
35	91	91.9	47340	9	AC069387
36	91	91.9	112085	9	AC022132
37	91	91.9	112085	9	AC026712
38	91	91.9	123182	9	AC093675
39	91	91.9	123253	9	AC078963
40	91	91.9	129770	9	AC104684
41	91	91.9	136901	9	AC073626
42	91	91.9	138003	2	AC018346
43	91	91.9	139111	9	AC008768
44	91	91.9	163166	9	AL356632
45	91	91.9	164943	9	AP005272

ALIGNMENTS

RESULT 1
LOCUS AX007980 10499 bp DNA Linear PAT 06-SEP-2000
DEFINITION Sequence 3 from Patent WO9967595.
ACCESSION AX007980
VERSION AX007980.1 GI:9995677
KEYWORDS
SOURCE
ORGANISM human.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
AUTHORS Perlin, J.P., Rieger, F. and Alliel, P.M.
TITLE Nucleic sequence and deduced protein sequence family with human
endogenous retroviral motifs, and their uses

Pred. No. is the number of results predicted by chance to have a

JOURNAL Patent: WO 9967395-A 3 29-DEC-1999;
INST NAT SANTE RECH MED (FR); PERIN JEAN PIERRE (FR); RIEGER
FRANCOIS (FR); ALLIEU PATRICK M (FR)
FEATURES location/Qualifiers
source 1.10499
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 3048 a 2676 c 2280 g 2495 t
ORIGIN

Query Match 100.0%; Score 99; DB 6; Length 10499;
Best Local Similarity 100.0%; Pred. No. 9.7e-24;
Matches 99; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCACCCCTGAGCGCAAGCAATGTTGGCATGCTGTAAGACACCTAGATCCAGCA 60
DB 2502 TCACCCCTGAGCGCAAGCAATGTTGGCATGCTGTAAGACACCTAGATCCAGCA 2561

QY 61 GCCCAGACCCCTTCTTGTGTCAGAAAGCGCGGAAA 99
DB 2562 GCCCAGACCCCTTCTTGTGTCAGAAAGCGCGGAAA 2600

RESULT 2
AX329572 56093 bp DNA linear PAT 09-JAN-2002
LOCUS
DEFINITION Sequence 81 from Patent WO0194629.
ACCESSION AX329572
VERSION AX329572.1 GI:18102550
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

REFERENCE 1
AUTHORS Young, P. E., Augustus, M., Carter, K. C., Ebner, R., Endress, G.,
Horrigan, S., Soppet, D. R. and Weaver, Z.
TITLE Cancer gene determination and therapeutic screening using signature
gene sets
JOURNAL Patent: WO 0194629-A 81 13-DEC-2001;
Avalon Pharmaceuticals (US)
FEATURES location/Qualifiers
source 1.56093
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 16164 a 12346 c 10702 g 16881 t
ORIGIN

Query Match 100.0%; Score 99; DB 6; Length 56093;
Best Local Similarity 100.0%; Pred. No. 1.1e-23;
Matches 99; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCACCCCTGAGCGCAAGCAATGTTGGCATGCTGTAAGACACCTAGATCCAGCA 60
DB 30502 TCACCCCTGAGCGCAAGCAATGTTGGCATGCTGTAAGACACCTAGATCCAGCA 30561

QY 61 GCCCAGACCCCTTCTTGTGTCAGAAAGCGCGGAAA 99
DB 30562 GCCCAGACCCCTTCTTGTGTCAGAAAGCGCGGAAA 30600

RESULT 3
HSAC000064 56093 bp DNA linear PRI 13-NOV-1996
LOCUS
DEFINITION Human BAC clone RG083M05, from 7q21-7q22, complete sequence.
ACCESSION AC000064
VERSION AC000064.1 GI:1669369
KEYWORDS HTG.
SOURCE Homo sapiens.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 56093)

AUTHORS Pauley, A.
TITLE The sequence of H. sapiens BAC clone RG083M05
JOURNAL Unpublished (1996)
REFERENCE 2 (bases 1 to 56093)
AUTHORS Waterston, R.
TITLE Direct Submission
JOURNAL Submitted (13-NOV-1996)
COMMENT Genome Sequencing Center
Department of Genetics, Washington University
St. Louis, MO 63108, USA
e-mail: sapiens@watson.wustl.edu

NOTICE: This sequence may not represent the entire insert of this
clone. It may be shorter because we only sequence overlapping
sections once, or longer because we provide a small overlap between
neighboring submissions.

This sequence was finished as follows unless otherwise noted:
all regions were double stranded or sequenced with an alternate
chemistry; an attempt was made to resolve all sequencing problems,
such as compressions and repeats; all regions were covered by
sequence from more than one subclone; and the assembly was
confirmed by restriction digest.

SOURCE INFORMATION:
This clone is from the first release of the human BAC library. The
library contains cloned DNA from a human male fibroblast cell line
978SK. For references see: Shizuya et al., Proc. Natl. Acad. Sci.
89:8794-8797 (1992); Kim et al., Genomics 34:213-218 (1996).
VECTOR: pBEO
Selection: chloramphenicol

NEIGHBORING SEQUENCE INFORMATION:
The orientation of this clone is unknown. Actual start of this
clone is at base position 1 of H_RG083M05; actual end is at 56093
of H_RG083M05

FEATURES
source 1.56093
location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="7"
/map="7q21-7q22"
/clone_lib="CITB-978SK-B"
/clone_1lb="CITB-978SK-B"
complement(838..1131)
/rpl_family="ALU"
/rpl_16971
/gene="WUGSC:H_RG083M05.1"
join<1360..1503,4181..4370,4587..4774,6422..6556,
9483..9547,11631..11773,11864..12021,13131..13256,
14885..14988,16349..16546,16837..16971)
/gene="WUGSC:H_RG083M05.1"
/note="Arpase; strong similarity to peroxisome
biosynthesis protein Pasi (PDB:1172019); coded for by
human cDNA C04279 (NID:g1467530)"
/codon_start=1
/protein_id="BAB46346.1"
/db_xref="GI:1669371"
/translation="KRLNIOKLTLEVAESEAVMOPSVLLDLDLIALAPVPEHEH
SPDARCEILCNVINKIKLDCDINKFTDDLDLQVAKETGFGVARDFTVYDAIHSRL
SROSISTREKLVLTDDFOKALRGFLPALRSVNLHKKPDLMDKRGVAGHEROIIMD
TIOIPAKVCLKREKYPELANIPIDRGCIILXGPGCGKTLIACVIAEESMNTSV
KPELISKTIGSEDAVRDIFTRDAQAKCITLFFDEFESIAIRRGHDTGTYDRVYNQ
LITOLDGVEGLGQVYVLAITSRDLIDPALLRGRIDKCYCPPPQVITVYLESTQ
QMHSLEIVSRLEILNVLSDSLADVDVQHAASVYDSFGADLALALYNAOLEALHG
MLSKMSLEIVPESEKFNRYLKEGSEYESELSDSISDCLASPSMTODLPCVP
GKDOLESOPPVLRITASOEGCOELTOEOROLRADISIIKGRYRSOGEDESMNGQPI
KTRLAISOSHLMATAGHTRPSISEDDMKNFABL"
complement(4948..5130)
/rpl_family="ALU"
complement(6581..7133)
repeat_region
repeat_region
repeat_region

```

repeat_region /rpt_family="L1"
complement(7767..8037)
repeat_region /rpt_family="ALU"
complement(8186..8472)
misc_feature /rpt_family="ALU"
8473..8625
/Note="match to human 3' EST H75782 (NID:g1049794), bases
287-444"
misc_feature 8841..9161
/Note="match to human 5' EST H75921 (NID:g1050050), bases
21-348"
misc_feature 9481..9547
/Note="WUGSC:H_RG083M05.1"
/Note="match to human 5' EST N2627 (NID:g1130501), bases
276-343"
repeat_region /rpt_family="ALU"
complement(12612..12907)
misc_feature 13670..13793
/Note="WUGSC:H_RG083M05.1"
/Note="match to human 5' EST H41382 (NID:g17434), bases
143-266"
repeat_region /rpt_family="ALU"
13794..13877
misc_feature 13878..13906
/Note="WUGSC:H_RG083M05.1"
/Note="match to human 5' EST H41382 (NID:g17434), bases
30-58"
repeat_region /rpt_family="ALU"
13907..14104
complement(14110..14137)
repeat_region /rpt_family="L1"
complement(15618..15907)
repeat_region /rpt_family="ALU"
17227..17522
misc_feature 18667..19235
/Note="match to human fetal brain 5' EST D61494
(NID:g970409), bases 1-255, and to human 3' EST R07476
19550..19670
/Note="WUGSC:H_RG083M05.1"
/Note="match to human 5' EST H41382 (NID:g17434), bases
21507..37303
/Note="similarity to various SS-RNA virus polyproteins;
pseudogene; region of matches and close matches to
multiple human ESTs, see R68740 (NID:g842257)"
37316..37489
/Note="Grail prediction, score = 80"
/Note="evidence-not-experimental"
complement(38938..39224)
repeat_region /rpt_family="ALU"
39225..39707
/Note="match to multiple human ESTs, see N03113
(NID:g1148633)"
39800..40085
repeat_region /rpt_family="ALU"
complement(40247..40538)
repeat_region /rpt_family="ALU"
complement(40632..40924)
repeat_region /rpt_family="ALU"
complement(42283..42891)
repeat_region /rpt_family="ALU"
complement(45474..45613)
misc_feature /rpt_family="ALU"
complement(45614..45737)
/Note="match to human 3' EST H48898 (NID:g988738), bases
129-333"
misc_feature complement(46107..47026)
/Note="match to multiple human ESTs, see N81064
(NID:g1243765), H48897 (NID:g988737), and M78831
(NID:g273146)"
repeat_region complement(47027..47318)

```

```

misc_feature /rpt_family="ALU"
complement(47365..47782)
/Note="match to multiple human ESTs, see W37495
(NID:g1319089)"
misc_feature 47898..48115
/Note="match to human 5' EST H62306 (NID:g1015138), bases
93-368"
repeat_region complement(48116..48405)
/Note="WUGSC:H_RG083M05.1"
/Note="match to human 5' EST N29952 (NID:g1148472), bases
290-455, and 5' EST R12730 (NID:g765806)"
misc_feature complement(48787..49405)
/Note="WUGSC:H_RG083M05.1"
/Note="match to human 3' EST R65794 (NID:g838432), bases
309-440"
repeat_region complement(49638..49672)
/Note="WUGSC:H_RG083M05.1"
/Note="match to human 3' EST N29952 (NID:g1148472) and 5'
EST N29938 (NID:g1148458), sequences are from opposite
ends of the same clone"
misc_feature complement(49698..51806)
/Note="WUGSC:H_RG083M05.1"
/Note="match to human 3' EST N29952 (NID:g1148472) and 5'
EST N29938 (NID:g1148458), sequences are from opposite
ends of the same clone"
gene complement(49698..51806)
/Note="WUGSC:H_RG083M05.1"
/Note="match to human 3' EST N29952 (NID:g1148472) and 5'
EST N29938 (NID:g1148458), sequences are from opposite
ends of the same clone"
CDS complement(49698..51806)
/Note="WUGSC:H_RG083M05.1"
/Note="match to human 3' EST N29952 (NID:g1148472) and 5'
EST N29938 (NID:g1148458), sequences are from opposite
ends of the same clone"
exon complement(51576..51758)
/Note="WUGSC:H_RG083M05.1"
/Note="match to human 3' EST N29952 (NID:g1148472) and 5'
EST N29938 (NID:g1148458), sequences are from opposite
ends of the same clone"
repeat_region complement(52052..52329)
/Note="WUGSC:H_RG083M05.1"
/Note="match to human 3' EST N29952 (NID:g1148472) and 5'
EST N29938 (NID:g1148458), sequences are from opposite
ends of the same clone"
misc_feature 55557..55843
/Note="match to human EST M79192 (NID:g273505) base 2-289"

```

Query Match 100.0%; Score 99; DB 9; Length 56093;
 Best Local Similarity 100.0%; Pred. No. 1.1e-23;
 Matches 99; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 TCACCCCTGAGCGAAGGCAATGTGGCATGCTGGTAAGGACCACTAGATCCAGCA 60
DB 30502 TCACCCCTGAGCGAAGGCAATGTGGCATGCTGGTAAGGACCACTAGATCCAGCA 30561
QY 61 GCCCAGACCCCTTCTTGTGTGTCAGAAAGCGCGGAAA 99
DB 30562 GCCCAGACCCCTTCTTGTGTGTCAGAAAGCGCGGAAA 30600

```

RESULT 4
 AC007566/c 149194 bp DNA linear PRI 01-MAR-2002
 LOCUS Homo sapiens BAC clone CTB-1065 from 7q21-7q22, complete sequence.
 DEFINITION AC007566
 ACCESSION AC007566
 VERSION AC007566.2 GI:11181861
 KEYWORDS HTG.
 SOURCE Homo sapiens.
 ORGANISM Homo sapiens.
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 149194)
 AUTHORS Sulston,J.E. and Waterston,R.
 TITLE Toward a complete human genome sequence
 JOURNAL Genome Res. 8 (11), 1097-1108 (1998)

MEDLINE
99063792
PUBMED
9847074
2 (bases 1 to 149194)
Du, Z.
TITLE
The sequence of Homo sapiens BAC clone CTB-10G5
JOURNAL
Unpublished (2001)
REFERENCE
3 (bases 1 to 149194)
WATERSTON, R.H.
TITLE
Direct Submission
JOURNAL
Submitted (15-MAY-1999) Genome Sequencing Center, Washington University School of Medicine, 4444 Forest Park Parkway, St. Louis, MO 63108, USA
REFERENCE
4 (bases 1 to 149194)
WATERSTON, R.
TITLE
Direct Submission
JOURNAL
Submitted (02-OCT-2000) Department of Genetics, Washington University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
REFERENCE
5 (bases 1 to 149194)
WATERSTON, R.H.
TITLE
Direct Submission
JOURNAL
Submitted (16-NOV-2000) Genome Sequencing Center, Washington University School of Medicine, 4444 Forest Park Parkway, St. Louis, MO 63108, USA
REFERENCE
6 (bases 1 to 149194)
WATERSTON, R.H.
TITLE
Direct Submission
JOURNAL
Submitted (03-JAN-2002) Genome Sequencing Center, Washington University School of Medicine, 4444 Forest Park Parkway, St. Louis, MO 63108, USA
REFERENCE
7 (bases 1 to 149194)
WATERSTON, R.H.
TITLE
Direct Submission
JOURNAL
Submitted (06-FEB-2002) Genome Sequencing Center, Washington University School of Medicine, 4444 Forest Park Parkway, St. Louis, MO 63108, USA
REFERENCE
8 (bases 1 to 149194)
WATERSTON, R.
TITLE
Direct Submission
JOURNAL
Submitted (01-MAR-2002) Department of Genetics, Washington University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
On Nov 16, 2000 this sequence version replaced gi:4835815.

Genome Center
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: <http://genome.wustl.edu/gsc>
Contact: saplens@wustl.wustl.edu

Summary Statistics
Center project name: H_RG010605

NOTICE: This sequence may not represent the entire insert of this clone. It may be shorter because we only sequence overlapping clone sections once, or longer because we provide a small overlap between neighboring data submissions.

This sequence was finished as follows unless otherwise noted:
all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

MAPPING INFORMATION:

The sequence of this clone was established as part of a mapping and sequencing collaboration between the NHGRI Chromosome 7 Mapping Project (Eric D. Green, Director), John D. McPherson in the Department of Genetics (Washington University), and the Washington University Genome Sequencing Center. For additional information about the map position of this sequence, see <http://www.nhgri.nih.gov/DIR/CTB/CHR7>, send <mailto:egreen@nhgri.nih.gov>, or see <http://genome.wustl.edu/gsc>

SOURCE INFORMATION:
Clone CTB-10G5 is from the first release of the human BAC library CTB-978SK-B. The library contains cloned DNA from the male fibroblast cell line 978SK. See: Shizuya et al., Proc. Natl. Acad. Sci. USA 89:8794-7 (1992); U-J. Kim et al., Genomics 34:213-8 (1996). This clone is available from Research Genetics, Inc. (<http://www.resgen.com>).
VECTOR: pBelobAC11
Selection: chloramphenicol

NEIGHBORING SEQUENCE INFORMATION:
The clone sequenced to the left is RP5-911H5, 200 base pair overlap. Actual start of this clone is at base position 195 of CTB-10G5; actual end is at base position 150532 of CTB-10G5.

FEATURES

source

repeat_region

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

Location/Qualifiers
1. 149194
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="7"
/map="7q21-7q22"
/clone="CTB-10G5"
/clone_11b="CTB-978SK-B"
1. 1634
/rpt_family="I1"
2248. 2388
/note="match to EST BG752883 (NID:g14063536)"
2248. 2387
/note="match to EST AA149693 (NID:g1720635) zn99d06.r1"
2248. 2387
/note="match to EST AW579261 (NID:g7254310)"
2248. 2387
/note="match to EST BG766882 (NID:g14077335)"
2248. 2387
/note="match to EST B1160365 (NID:g14620366)"
2248. 2374
/note="match to Homo sapiens EST BF758865 (NID:g12106765)"
2248. 2287
/note="match to EST AV686676 (NID:g10288539)"
2253. 2387
/note="match to EST AU123510 (NID:g10948226)"
2344. 2387
/note="match to EST AV686676 (NID:g10288539)"
2696. 3066
/rpt_family="MaLR"
3108. 3392
/rpt_family="Alu"
3540. 3628
/note="match to EST AA425526 (NID:g2106267) zw48b03.r1"
3540. 3628
/note="similar to Mus musculus EST BB253526 (NID:g8946272)"
3542. 3632
/note="match to EST AV686676 (NID:g10288539)"
3542. 3628
/note="match to EST AA149693 (NID:g1720635) zn99d06.r1"
3542. 3628
/note="match to EST AU123510 (NID:g10948226)"
3542. 3628
/note="match to EST BG766882 (NID:g14077335)"
3542. 3628
/note="match to EST B1160365 (NID:g14620366)"
3542. 3609
/note="match to EST AW579261 (NID:g7254310)"
3562. 3628
/note="match to EST BE272564 (NID:g9146913)"
3598. 3628
/note="match to EST BE299708 (NID:g9183456)"
3714. 3785
/note="match to EST BG260659 (NID:g12770475)"
3717. 3785

/note="similar to Mus musculus EST BE994936
(NID:g10678674)"
misc_feature 3719..3785
/note="match to EST AA149693 (NID:g1120635) zn99d06.r1"
misc_feature 3719..3785
/note="match to EST AA425526 (NID:g2106267) zw48b03.r1"
misc_feature 3719..3785
/note="match to EST AU123510 (NID:g10948226)"
misc_feature 3719..3785
/note="match to EST BE272564 (NID:g9146913)"
misc_feature 3719..3785
/note="match to EST BE299708 (NID:g9183456)"
misc_feature 3719..3785
/note="match to EST BG766882 (NID:g14077535)"
misc_feature 3719..3785
/note="match to EST BI160365 (NID:g14620366)"
misc_feature 3719..3785
/note="similar to Mus musculus EST BB253526
(NID:g8946272)"
misc_feature 3860..3862
/note="match to EST AV686676 (NID:g10288539)"
repeat_region 4730..4827
/rpt_family="L2"
repeat_region 4834..4925
/rpt_family="L1"
misc_feature 5715..5860
/note="match to EST AA425526 (NID:g2106267) zw48b03.r1"
misc_feature 5715..5860
/note="match to EST AU123510 (NID:g10948226)"
misc_feature 5715..5860
/note="match to EST BE272564 (NID:g9146913)"
misc_feature 5715..5860
/note="match to EST BE299708 (NID:g9183456)"
misc_feature 5715..5860
/note="match to EST BG260659 (NID:g12770475)"
misc_feature 5715..5860
/note="match to EST BG766882 (NID:g14077535)"
misc_feature 5715..5860
/note="match to EST BI160365 (NID:g14620366)"
misc_feature 5715..5860
/note="similar to Mus musculus EST BE994936"

Query Match 100.0%; Score 99; DB 9; Length 149194;
Best Local Similarity 100.0%; Pred. No. 1,1e-23;
Matches 99; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 TCACCCCTGAGCGCAAGGCAATGTTGGCATGCTGTAAAGGACCACTAGATCCAGCA 60
|||||
Db 91502 TCACCCCTGAGCGCAAGGCAATGTTGGCATGCTGTAAAGGACCACTAGATCCAGCA 91443
|||||

OY 61 GCCCAGACCCCTTCTTGTGTGTCAGAAAGCGGGGAAA 99
|||||
Db 91442 GCCCAGACCCCTTCTTGTGTGTCAGAAAGCGGGGAAA 91404
|||||

RESULT 5
AF045450/c 40205 bp DNA linear PRI 20-MAR-1998
LOCUS Homo sapiens chromosome 21q22.3 cosmid Q11M15, complete sequence.
ACCESSION AF045450
VERSION AF045450.1 GI:2895783
KEYWORDS HTG.
SOURCE Homo sapiens.
ORGANISM Homo sapiens.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 40205)
AUTHORS Taudien,S., and Rosenthal,A.
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 40205)
AUTHORS Taudien,S., Nordstiek,G., Dagand,E., Hildmann,T., Drescher,B.,
Weber,J., Rosenthal,A. and Yaspo,M.L.
TITLE Direct Submission

JOURNAL Submitted (29-JAN-1998) Genome Analysis, Institut for Molecular
Biotechnology, Beutenbergstrasse 11, Jena 07745, Germany
FEATURES
Source
1..40205
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="21"
/map="21q22.3"
/clone="cosmid Q11M15"
complement(293..586)
/evidence-not_experimental
/rpt_family="AluSc"
complement(643..964)
/evidence-not_experimental
/rpt_family="MERVL"
complement(1053..1176)
/note="Genscan, score = 5.19%, comment = Internal-exon 124
bp frame: 2 phase: 1"
/evidence-not_experimental
complement(1251..1859)
/evidence-not_experimental
/rpt_family="HERV1"
complement(2296..3220)
/evidence-not_experimental
/rpt_family="HERV16"
2818..2889
/note="Xpound exon prediction, score = 82% (0%)"
/evidence-not_experimental
complement(3221..3658)
/evidence-not_experimental
/rpt_family="MLT1C"
3359..3395
/note="Xpound exon prediction, score = 84% (0%)"
/evidence-not_experimental
complement(3874..4176)
/evidence-not_experimental
/rpt_family="Aluub"
complement(4330..4433)
/evidence-not_experimental
/rpt_family="L2"
4585..4638
/note="GRAIL, score = 95.000%, comment = excellent shadow"
/evidence-not_experimental
complement(5113..7899)
/evidence-not_experimental
/rpt_family="L1PAL4"
complement(6119..6598)
/note="GRAIL, score = 63.000%, comment = good"
/evidence-not_experimental
complement(7927..8045)
/evidence-not_experimental
/rpt_family="FLAV_A"
8048..8178
/evidence-not_experimental
/rpt_family="L2"
8771..8876
/note="WZEF, score = 67.4%"
/evidence-not_experimental
complement(8809..9187)
/evidence-not_experimental
/rpt_family="MER1A"
8837..8876
/note="GRAIL, score = 55.000%, comment = good shadow"
/evidence-not_experimental
complement(9113..9213)
/note="GRAIL, score = 55.000%, comment = good"
/evidence-not_experimental
complement(9190..9258)
/evidence-not_experimental
/rpt_family="L1MC3"
9259..9647
/evidence-not_experimental
/rpt_family="L1MC3"

```
repeat_region complement(10138..10427)
/evidence-not_experimental
/rpt_family="MLT2E"
repeat_region 10662..10875
/evidence-not_experimental
/rpt_family="MER58A"
exon complement(11072..11156)
/note="GRAIL, score = 86.000%, comment = excellent"
/evidence-not_experimental
complement(11072..11129)
/note="Xpound exon prediction, score = 89% (0%)"
exon 11478..11597
/evidence-not_experimental
/rpt_family="L2"
repeat_region 11570..11704
/note="MZF, score = 66.1%"
/evidence-not_experimental
exon 11570..11704
/note="GRAIL, score = 99.000%, comment = excellent shadow"
/evidence-not_experimental
complement(11687..11735)
repeat_region 11774..12320
/evidence-not_experimental
/rpt_family="MIR"
repeat_region 11774..12320
/evidence-not_experimental
/rpt_family="MLT1E"
exon complement(11828..11949)
/note="GRAIL, score = 50.000%, comment = good"
/evidence-not_experimental
12623..12843
repeat_region 13102..13457
/evidence-not_experimental
/rpt_family="AluXs"
repeat_region 13102..13457
/evidence-not_experimental
/rpt_family="THB1B"
exon complement(13481..13519)
/note="GRAIL, score = 65.000%, comment = good"
/evidence-not_experimental
15065..15188
repeat_region 15065..15188
/evidence-not_experimental
/rpt_family="HERVL"
misc_feature 15130..15378
/note="GC score = 10.60 (249bp); Region: GC content"
/evidence-not_experimental
15369..16268
repeat_region 16268..16730
/evidence-not_experimental
/rpt_family="MERVL"
exon complement(16623..16714)
/note="GRAIL, score = 72.000%, comment = good"
/evidence-not_experimental
16661..16730
/note="MZF, score = 93.5%"
exon 16880..17318
/evidence-not_experimental
repeat_region 17326..17436
/evidence-not_experimental
/rpt_family="MLT1C"
repeat_region 17326..17436
/evidence-not_experimental
/rpt_family="LTRL6C"
repeat_region 17540..17675
/evidence-not_experimental
/rpt_family="MIR"
repeat_region 17805..17864
/evidence-not_experimental
/rpt_family="MLT1C"
exon 18076..18098
/note="GRAIL, score = 45.000%, comment = marginal shadow"
/evidence-not_experimental
complement(18219..18349)
exon 18349..18499
/note="Genscan, score = 2.99%, comment = Internal_exon 131
bp frame: 1 phase: 2"
/evidence-not_experimental
```

```
repeat_region complement(18717..19211)
/evidence-not_experimental
/rpt_family="MLT1C"
exon complement(19252..19399)
/note="GRAIL, score = 71.000%, comment = good"
/evidence-not_experimental
19269..19350
exon 19269..19350
/note="GRAIL, score = 81.000%, comment = excellent shadow"
/evidence-not_experimental
complement(19379..19399)
exon 19379..19399
/note="Xpound exon prediction, score = 60% (0%)"
/evidence-not_experimental
complement(19501..19881)
exon 19501..19881
/note="MZF, score = 50.6%"
/evidence-not_experimental
19945..20020
repeat_region 19945..20020
/note="homology = 100.00%, score = 38, counts = 2"
/evidence-not_experimental
/rpt_type=tandem
/rpt_unit=ttcttgcacacctctctgtgagagtcgattatcgc
repeat_region 20223..20643
/evidence-not_experimental
/rpt_family="MLT1E"
repeat_region 20645..20910
/evidence-not_experimental
/rpt_family="MLT1F"
repeat_region 20961..21007
/evidence-not_experimental
/rpt_family="MLT1F"
repeat_region 21483..21781
/evidence-not_experimental
/rpt_family="AluJb"
exon complement(21504..21756)
/note="MZF, score = 79.4%"
/evidence-not_experimental
complement(22325..22806)
repeat_region 22325..22806
/evidence-not_experimental
/rpt_family="MLT1C"

Query Match 96.8%; Score 95.8; DB 9; Length 40205;
Best Local Similarity 98.0%; Pred. No. 1.4e-22;
Matches 97; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 TCACCCCTGAGCGCAAGGCATGTTGGCGATCGTGAAGGACCCACTAGATCCAGCA 60
DB 32980 TCACCCCTGAGCGCAAGGCATGTTGGCGAGCGTGAAGGACCCACTAGATCCAGCA 32921
QY 61 GCCCAGACCCCTTCTTTGTGTCAGAAAGCGCGGAAA 99
DB 32920 GCCCGAGACCCCTTCTTTGTGTCAGAAAGCGCGGAAA 32882

RESULT 6
AC027500 183951 bp DNA linear HTG 27-APR-2000
LOCUS Homo sapiens chromosome 3 clone RP11-705F24 map 3, WORKING DRAFT
DEFINITION AC027500
ACCESSION AC027500
VERSION AC027500.2 GI:7651959
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE Homo sapiens.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 183951)
AUTHORS Birren,B., Linton,L., Nusbaum,C. and Lander,E.
TITLE Homo sapiens chromosome 3, clone RP11-705F24
JOURNAL Unpublished
RECORD 2 (bases 1 to 183951)
AUTHORS Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
Anderson,S., Baldwin,J., Barna,N., Bastien,V., Beda,F.,
Boguslavsky,L., Bouknight,B., Brown,A., Burkett,G.,
Campopiano,A., Castle,A., Choepel,Y., Colangelo,M., Collins,S.,
```

TITLE
JOURNAL
COMMENT

Collamore, A., Cooke, P., Dearellano, K., Dewar, K., Diaz, J. S.,
Dodge, S., Domingo, M., Doyle, M., Ferreira, P., Fitzhugh, W., Gage, D.,
Gadigan, J., Gardyna, S., Ginde, S., Goyette, M., Graham, L.,
Grand-Pierre, N., Grant, G., Hegos, B., Hearford, A., Horton, L.,
Howland, J. C., Iliev, I., Johnson, R., Jones, C., Kann, L., Karakas, A.,
Klein, J., Lacroque, K., Lamazares, R., Landers, T., Lehoczy, J.,
Levine, R., Lieu, C., Liu, G., Locke, K., Macdonald, P., Marquis, N.,
McCarthy, M., McKean, P., McGurk, A., McKernan, K., McPheters, R.,
Meldrum, J., Menus, L., Minova, T., Miranda, C., Mlenga, V., Morrow, J.,
Murphy, T., Naylor, J., Norman, C. H., O'Connor, T., O'Donnell, P.,
O'Neill, D., Oliver, T. M., Oliver, J., Peterson, K., Pierre, N.,
Pisani, C., Pollara, V., Raymond, C., Riley, R., Rogov, P., Rothman, D.,
Roy, A., Santos, R., Schauer, S., Severy, P., Spencer, B.,
Stange-Thomann, N., Stojanovic, N., Subramanian, A., Talamas, J.,
Teafaye, S., Theodore, J., Tirrell, A., Travers, M., Triggillo, J.,
Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W. J.,
Young, G., Zainoun, J., Zimmer, A. and Zody, M.

Direct Submission

Submitted (30-MAR-2000) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Apr 27, 2000 this sequence version replaced gi:7342244.
All repeats were identified using RepeatMasker:

Smit, A.F.A. & Green, P. (1996-1997)
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

Genome Center

Center: Whitehead Institute/ MIT Center for Genome Research

Center code: WIBR

Web site: <http://www-seq.wi.mit.edu>

Contact: sequence_submissions@genome.wi.mit.edu

Project Information

Center project name: L8456

Center clone name: 705_F_24

Summary Statistics

Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.960731
Consensus quality: 175817 bases at least Q40
Consensus quality: 180619 bases at least Q30
Consensus quality: 181937 bases at least Q20
Insert size: 18500; agarose-fp
Insert size: 182551; sum-of-contigs
Quality coverage: 4.9 in Q20 bases; agarose-fp
Quality coverage: 4.9 in Q20 bases; sum-of-contigs

* NOTE: This is a 'working draft' sequence. It currently
* consists of 15 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

```

* 1. 1062: contig of 1062 bp in length
* 1063 1162: gap of 100 bp
* 1163 2742: contig of 1580 bp in length
* 2743 2842: gap of 100 bp
* 2843 4245: contig of 1403 bp in length
* 4246 4345: gap of 100 bp
* 4346 6301: contig of 1956 bp in length
* 6302 6401: gap of 100 bp
* 6402 9026: contig of 2625 bp in length
* 9027 9126: gap of 100 bp
* 9127 13381: contig of 4255 bp in length
* 13382 13481: gap of 100 bp
* 13482 19288: contig of 5807 bp in length
* 19289 19388: gap of 100 bp
* 19389 25973: contig of 6585 bp in length
* 25974 26073: gap of 100 bp
* 26074 35483: contig of 9416 bp in length
* 35490 35589: gap of 100 bp
* 35590 46320: contig of 10631 bp in length
* 46321 46320: gap of 100 bp
* 46321 62888: contig of 16568 bp in length

```

```

* 62889 62988: gap of 100 bp
* 62989 81495: contig of 18507 bp in length
* 81496 81595: gap of 100 bp
* 81596 99524: contig of 17923 bp in length
* 99525 99624: gap of 100 bp
* 99625 129294: contig of 29670 bp in length
* 129295 129394: gap of 100 bp
* 129395 183951: contig of 54557 bp in length.

```

FEATURES
Location/Qualifiers

Source

```

1. 183951
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="3"
/map="3"
/clone="RP11-705P24"
/clone_lib="RP11 Human Male BAC"
1. 1062

```

misc_feature

```

/note="assembly_fragment"
1163. 2742

```

misc_feature

```

/note="assembly_fragment"
2843. 4245

```

misc_feature

```

/note="assembly_fragment"
4346. 6301

```

misc_feature

```

/note="assembly_fragment"
6402. 9026

```

misc_feature

```

/note="assembly_fragment"
9127. 13381

```

misc_feature

```

/note="assembly_fragment"
13482. 19288

```

misc_feature

```

/note="assembly_fragment"
clone_end:T7
vector_side:right"
19389. 25973

```

misc_feature

```

/note="assembly_fragment"
26074. 35489

```

misc_feature

```

/note="assembly_fragment"
clone_end:SP6
vector_side:left"
35590. 46320

```

misc_feature

```

/note="assembly_fragment"
46321. 62888

```

misc_feature

```

/note="assembly_fragment"
62989. 81495

```

misc_feature

```

/note="assembly_fragment"
81596. 99524

```

misc_feature

```

/note="assembly_fragment"
99625. 129294

```

misc_feature

```

/note="assembly_fragment"
129395. 183951

```

misc_feature

```

/note="assembly_fragment"

```

BASE COUNT

```

58152 a 35325 c 34804 g 54268 t 1402 others

```

ORIGIN

```

Query Match 96.8%; Score 95.8; DB 2; Length 183951;
Best Local Similarity 98.0%; Pred. No.1.5e-22;
Matches 97; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

```

QY 1 TCACCCCTGAGCGCAAGGCAATGTGGCATGCTGCTGTAAGGACCACTAGATCCAGCA 60
DB 30861 TCACCCCTGAGCGCAAGGCAATGTGGCATGCTGCTGTAAGGACCACTAGATCCAGCA 30920
QY 61 GCCCAGACCCCTTCTTGTGTGTCAGAAAGCGGGGAAA 99
DB 30921 GCCCAGACCCCTTCTTGTGTGTCAGAAAGCGGGGAAA 30959

```

RESULT 7

```

AC117444/c 104853 bp DNA linear PRI 31-JUL-2002
LOCUS Homo sapiens 3 BAC RP11-215L17 (Roswell Park Cancer Institute Human
DEFINITION BAC Library) complete sequence.
AC117444
AC117444.6 GI:22024337

```

KEYWORDS HTG.
SOURCE human.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
1 (bases 1 to 104853)

TITLE Direct Submission
JOURNAL Unpublished
AUTHORS Worley, K.C.
REFERENCE Submitted (10-APR-2002) Human Genome Sequencing Center, Department
JOURNAL of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 104853)
TITLE Direct Submission
JOURNAL Submitted (24-JUL-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
4 (bases 1 to 104853)
REFERENCE Submitted (31-JUL-2002) Human Genome Sequencing Center, Department
AUTHORS of Molecular and Human Genetics, Baylor College of Medicine, One
TITLE Baylor Plaza, Houston, TX 77030, USA
JOURNAL On Jul 31, 2002 this sequence version replaced gi:21908338.
COMMENT INFORMATION: <http://www.hgsc.bcm.tmc.edu/> or email
gc-helpebcm.tmc.edu

CLONE LENGTH: This sequence does not necessarily represent the
entire insert of this clone. Overlapping regions of clones are only

sequenced and submitted once, so the sequence for the remainder of
the insert may be found in the record for the adjacent clones.
Overlapping clones are noted at the beginning and end of the
features listing.

ANNOTATION OF FEATURES:

STSs are identified using ePCR (Genome Res. 7:541-550) searches
of a local database that includes entries from dbSTS, GDB, and
local mapping efforts.
Repeats are identified using RepeatMasker (A. Smit and P. Green,
unpublished.) for Human and Mouse sequences.
Genes and Region of sequence similarity are identified by BLAST
(Nuc. Acids Res. 25:3389-3402) similarity (expect < 1e-34) to the
EST and cDNA sequences. Genes demonstrate at least two exons
flanked by consensus splice sites that maintained sequence
continuity across the splice junctions. Sequences that are not
identical matches are annotated as similar.

SEQUENCING READ COVERAGE: Sequencing is completed to a minimum
standard of double strand coverage with a minimum of 2 clones and 2
reads with no ambiguities or 2 chemistries with a minimum of 2
clones and 3 reads with no ambiguities. If the sequence quality for
a region does not meet this standard, it will be indicated in the
annotation as Low Coverage.

QUALITY OF INDIVIDUAL BASES: This sequence meets stringent quality
standards - estimated error rate less than 1 per 10,000 bases.
Reports of lowest quality individual bases and measures of base
quality are listed below. Description of the metrics can be found
at URL: http://www.hgsc.bcm.tmc.edu:8088/quality/info/genbank_annotation.html.

FEATURES

SOURCE

1. 104853
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="3"
/clone="RP11-215L17"

MISC_FEATURE

1. 2000
/note="overlaps bases 154275..156277 of clone AC107015"
/function="clone overlap"

REPEAT_REGION

1. 954
/rpt_family="AT_n"

REPEAT_REGION

1576..1598
/rpt_family="AT_rich"

REPEAT_REGION

complement(2562..3082)
/rpt_family="MLT1G"

REPEAT_REGION

4553..4922
/rpt_family="MLT1A1"

REPEAT_REGION

5010..5035
/rpt_family="AT_rich"

REPEAT_REGION

5079..5406
/rpt_family="L2"

REPEAT_REGION

6323..6351
/rpt_family="(A)n"

REPEAT_REGION

6948..7099
/rpt_family="MIR"

REPEAT_REGION

7338..7447
/rpt_family="MER94"

REPEAT_REGION

8267..8294
/rpt_family="AT_rich"

REPEAT_REGION

9151..9242
/rpt_family="MLT1J"

REPEAT_REGION

9286..9509
/rpt_family="MLT1J2"

REPEAT_REGION

10028..10091
/rpt_family="(CATARA)n"

REPEAT_REGION

complement(13636..14644)
/rpt_family="MER45B"

REPEAT_REGION

15307..15335
/rpt_family="AT_rich"

REPEAT_REGION

15593..15628


```
/note="MZEf, score = 95.1%"
/evidence-not_experimental
6746. .6751
/note="Genscan, score = 1.83%, comment = Initial_exon 6 bp
frame: 1 phase: 0"
/evidence-not_experimental
6901. .7339
/evidence-not_experimental
/rpt_family="MLT1C"
7347. .7457
/evidence-not_experimental
/rpt_family="LIR16C"
complement(7561. .7696)
/evidence-not_experimental
/rpt_family="MIR"
complement(7819. .8220)
/evidence-not_experimental
/rpt_family="MLT1J"
8037. .8119
/note="GRAIL, score = 45.000%, comment = marginal"
/evidence-not_experimental
complement(8738. .9232)
/evidence-not_experimental
/rpt_family="MLT1C"
complement(9273. .9420)
/note="GRAIL, score = 71.000%, comment = good"
/evidence-not_experimental
9290. .9371
/note="GRAIL, score = 81.000%, comment = excellent shadow"
/evidence-not_experimental
9666. .10041
/note="homology = 100.00%, score = 38, counts = 2"
/evidence-not_experimental
/rpt_type=tandem
/rpt_unit=ctcttcacactctctgttgagaggtcggttatcgc
10244. 10664
/evidence-not_experimental
/rpt_family="MLT1E"
10665. .11028
/evidence-not_experimental
/rpt_family="MLT1F"
11318. .11417
/evidence-not_experimental
/rpt_family="LIMC/D"
11503. .11818
/evidence-not_experimental
/rpt_family="AluSg"
complement(11525. .11777)
/note="MZEf, score = 79.4%"
/evidence-not_experimental
complement(12346. .12827)
/evidence-not_experimental
/rpt_family="MLT1C"
12902. .13180
/evidence-not_experimental
/rpt_family="AluSg"
complement(13057. .13161)
/note="MZEf, score = 76.1%"
/evidence-not_experimental
complement(13226. .13364)
/evidence-not_experimental
/rpt_family="LIMC2"
complement(13613. .13908)
/evidence-not_experimental
/rpt_family="AluSx"
complement(14317. .14614)
/evidence-not_experimental
/rpt_family="LIR37A"
complement(14962. .15036)
/note="MZEf, score = 93%"
/evidence-not_experimental
15235. .15299
/evidence-not_experimental
```

```
/rpt_family="LIRP6"
complement(15469. .15497)
/note="Xpound exon prediction, score = 62% (0%)"
/evidence-not_experimental
16045. .16768
/evidence-not_experimental
/rpt_family="MER21A"
complement(16555. .16655)
/note="GRAIL, score = 54.000%, comment = good"
/evidence-not_experimental
16823. .16850
/evidence-not_experimental
/rpt_family="MER34"
complement(16851. .17208)
/evidence-not_experimental
/rpt_family="THB1A"
17209. .17632
/evidence-not_experimental
/rpt_family="MER34"
complement(17633. .17969)
/evidence-not_experimental
/rpt_family="AluSg1"
complement(18004. .18334)
/evidence-not_experimental
/rpt_family="LIR17"
complement(18335. .18632)
/evidence-not_experimental
/rpt_family="HERV17"
18352. .18420
/note="GRAIL, score = 51.000%, comment = good shadow"
/evidence-not_experimental
complement(18397. .18584)
/note="GRAIL, score = 66.000%, comment = good"
/evidence-not_experimental
complement(18643. .24647)
/evidence-not_experimental
/rpt_family="HERV17"
complement(19228. .19361)
/note="MZEf, score = 92.8%"
/evidence-not_experimental
20900. .21135
/note="MZEf, score = 51.4%"
/evidence-not_experimental
21150. .21275
/note="Xpound exon prediction, score = 74% (0%)"
/evidence-not_experimental
```

Query Match 95.2%; Score 94.2; DB 9; Length 142742;
Best Local Similarity 97.0%; Pred. No. 5.3e-22;
Matches 96; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 TCACCCCTGAGCCCAAGGCAATGTTGGCATGCTGTAAAGGACCACTGAAATCCAGCA 60
DB 22987 TCACCTCTGAGCCCAAGGCAATGTTGGCATGCTGTAAAGGACCACTGAAATCCAGCA 22928

QY 61 GCCCAGACCCCTTCTTGTGTCAAGAAAGCGGGGAAA 99
DB 22927 GCCCGAGCCCTTCTTGTGTCAAGAAAGCGGGGAAA 22889

RESULT 9
AC080070 164988 bp DNA 11near HTG 05-NOV-2000
LOCUS Homo sapiens chromosome 4 clone RP11-162D2, WORKING DRAFT SEQUENCE,
DEFINITION 24 unorderd pieces.
ACCESSION AC080070
VERSION AC080070.2 GI:11079592
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE Homo sapiens.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 164988)

AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNALWaterston,R.H.
The sequence of Homo sapiens clone
Unpublished
2 (bases 1 to 164988)Waterston,R.H.
Direct SubmissionSubmitted (23-SEP-2000) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
On Nov 3, 2000 this sequence version replaced gi:10280941.

COMMENT

----- Genome Center -----
Center: Washington University Genome Sequencing Center
Center code: WTGSC
Web site: http://genome.wustl.edu/gsc/index.shtml
----- Project Information -----
Center project name: H_NH0162D2
----- Summary Statistics -----
Sequencing vector: plasmid; 100%
Chemistry: Dye-Primer ET; 100% of reads
Chemistry: Dye-terminator Big Dye; 0% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 150371 bases at least Q40
Consensus quality: 155395 bases at least Q30
Consensus quality: 157776 bases at least Q20
Insert size: 16000; agarose-fp
Insert size: 162688; sum-of-contigs
Quality coverage: 4.08 in Q20 bases; agarose-fp
Quality coverage: 3.98 in Q20 bases; sum-of-contigs

* NOTE: This is a 'working draft' sequence. It currently
* consists of 24 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
1
* 3018: contig of 3018 bp in length
* 3019 3118: gap of unknown length
* 3119 6167: contig of 3049 bp in length
* 6168 6267: gap of unknown length
* 6268 9294: contig of 3027 bp in length
* 9295 9395: gap of unknown length
* 9395 12600: contig of 3206 bp in length
* 12601 12700: gap of unknown length
* 12701 16145: contig of 3444 bp in length
* 16145 16244: gap of unknown length
* 16245 20910: contig of 4666 bp in length
* 20911 25128: gap of unknown length
* 25129 25229: contig of 4119 bp in length
* 25230 30373: gap of unknown length
* 30374 30473: gap of unknown length
* 30474 34806: contig of 4333 bp in length
* 34807 34906: gap of unknown length
* 34907 39962: contig of 5056 bp in length
* 39963 40062: gap of unknown length
* 40063 46152: contig of 6090 bp in length
* 46153 46252: gap of unknown length
* 46253 52245: contig of 5993 bp in length
* 52246 52345: gap of unknown length
* 52346 58622: contig of 6277 bp in length
* 58623 58722: gap of unknown length
* 58723 65105: contig of 6383 bp in length
* 65106 72475: gap of unknown length
* 72476 72575: contig of 7270 bp in length
* 72576 80000: gap of unknown length
* 80001 80100: contig of 7425 bp in length
* 80101 90367: gap of unknown length
* 90368 90467: contig of 10267 bp in length
* 90467: gap of unknown length

FEATURES

Source

* 90468 99743: contig of 9276 bp in length
* 99744 99843: gap of unknown length
* 99844 114077: contig of 14234 bp in length
* 114078 114177: gap of unknown length
* 114178 128059: contig of 13882 bp in length
* 128060 128159: gap of unknown length
* 128160 144663: contig of 16504 bp in length
* 144664 144764: gap of unknown length
* 144765 160610: contig of 15847 bp in length
* 160611 160710: gap of unknown length
* 160711 162735: contig of 2025 bp in length
* 162736 162835: gap of unknown length
* 162836 164988: contig of 2153 bp in length.
Location/Qualifiers
1. 164988
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="4"
/clone="RP11-162D2"
1. 3018
/note="assembly_name:Contig10"
3119. 6167
/note="assembly_name:Contig11"
6268. 9294
/note="assembly_name:Contig12"
9395. 12600
/note="assembly_name:Contig13"
12701. 16144
/note="assembly_name:Contig14
clone_end:T7
vector_side:right"
16245. 20910
/note="assembly_name:Contig15"
21011. 25129
/note="assembly_name:Contig16"
25230. 30373
/note="assembly_name:Contig17"
30474. 34806
/note="assembly_name:Contig18"
34907. 39962
/note="assembly_name:Contig19"
40063. 46152
/note="assembly_name:Contig20"
46253. 52245
/note="assembly_name:Contig21"
52346. 58622
/note="assembly_name:Contig22"
58723. 65105
/note="assembly_name:Contig23"
65206. 72475
/note="assembly_name:Contig24"
72576. 80000
/note="assembly_name:Contig25"
80101. 90367
/note="assembly_name:Contig26"
90468. 99743
/note="assembly_name:Contig27"
99844. 114077
/note="assembly_name:Contig28"
114178. 128059
/note="assembly_name:Contig29"
128160. 144663
/note="assembly_name:Contig30"
144764. 160610
/note="assembly_name:Contig31"
160711. 162735
/note="assembly_name:Contig8"
162836. 164988
/note="assembly_name:Contig9"

BASE COUNT 48065 a 33578 c 34317 g 46717 t 2311 others
ORIGIN

Query Match

95.2%; Score 94.2; DB 2; Length 164988;

Best Local Similarity 97.0%; Pred. No. 5.4e-22;
Matches 96; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 TCACCCCTGACCAAGCAATGTTGGCATCTGTAAAGACCACTAGATCCAGCA 60
|||||
Db 70146 TCACCCCTGACCAAGCAATGTTGGCATCTGTAAAGACCACTAGATCCAGCA 70205
|||||

QY 61 GCCCAGACCCCTTCTTGTGTGTCAGAAAGCGGAGAA 99
|||||
Db 70206 GCCCAGACCCCTTCTTGTGTGTCAGAAAGCGGAGAA 70244
|||||

RESULT 10
AC073066 169286 bp DNA linear PRI 09-JAN-2002
LOCUS Homo sapiens BAC clone RP11-149P12 from 2, complete sequence.
DEFINITION AC073066
AC073066
AC073066.5 GI:15778739
KEYWORDS HTG.
SOURCE Homo sapiens.
ORGANISM Homo sapiens.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 169286)
Sulston, J.E. and Waterston, R.
Toward a complete human genome sequence
Genome Res. 8 (11), 1097-1108 (1998)
99063792
PUBMED 9847074

REFERENCE 2 (bases 1 to 169286)
Desai, A., Cordum, H., Maupin, R. and Elliott, G.
The sequence of Homo sapiens BAC clone RP11-149P12
Unpublished (2001)
3 (bases 1 to 169286)
Waterston, R.H.
Direct Submission
Submitted (08-JUN-2000) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
4 (bases 1 to 169286)
Waterston, R.H.
Direct Submission
Submitted (26-SEP-2001) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
5 (bases 1 to 169286)
Waterston, R.
Direct Submission
Submitted (09-JAN-2002) Department of Genetics, Washington
University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
On Sep 26, 2001 this sequence version replaced gi:15528927.
----- Genome Center
Center: Washington University Genome Sequencing Center
Center code: MUGSC
Web site: <http://genome.wustl.edu/gsc>
Contact: sapiens@wustl.edu
----- Summary Statistics
Center project name: H_NH0149P12

COMMENT

NOTICE: This sequence may not represent the entire insert of this clone. It may be shorter because we only sequence overlapping clone sections once, or longer because we provide a small overlap between neighboring data submissions.

This sequence was finished as follows unless otherwise noted:
all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

MAPPING INFORMATION:
Mapping information for this clone was provided by Dr. John D. McPherson, Department of Genetics, Washington University, St. Louis MO. For additional information about the map position of this sequence, see <http://genome.wustl.edu/gsc>

SOURCE INFORMATION:
The RPc1-11 human BAC library was made from the blood of one male donor, as described by Osoegawa, K., Woon, P. Y., Zhao, B., Frengen, E., Tatenno, M., Catanesse, J. J. and de Jong, P. J. (1998) An improved approach for construction of bacterial artificial chromosome libraries. Genomics 51:1-8. The clone may be obtained either from Research Genetics, Inc. (<http://www.resgen.com>) or Pieter de Jong and coworkers at the Roswell Park Cancer Institute (<http://bacpac.med.buffalo.edu>)
VECTOR: pBAC3.6

NEIGHBORING SEQUENCE INFORMATION:
Actual start of this clone is at base position 1 of RP11-149P12; actual end is at base position 169286 of RP11-149P12.

Data from AC024424 was used to finish this clone, AC073066.
Polymorphisms have been identified between AC073066 and AC024424.

FEATURES

source

1..169286
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="2"
/map="2"
/clone="RP11-149P12"
/clone_lib="RPc1-11"
15..103
/rpt_family="L1"
117..421
/rpt_family="L1"
658..690
/rpt_family="AT-rich"
1275..1321
/rpt_family="AT-rich"
1401..1786
/rpt_family="MELR"
2093..2122
/rpt_family="(TTTTC)n"
2160..2192
/rpt_family="AT-rich"
2174..2293
/rpt_family="L1"
2832..2893
/rpt_family="AT-rich"
3235..3373
/rpt_family="L2"
3597..3850
/rpt_family="Alu"
3630..3662
/rpt_family="(A)n"
3818..3879
/rpt_family="A-rich"
4072..4489
/rpt_family="MER2_type"
5015..5393
/rpt_family="MELR"
5520..5640
/rpt_family="L2"
5700..6149
/rpt_family="ERV1"
6185..6525
/rpt_family="ERV1"
6526..7677
/rpt_family="L1"
7662..7693
/rpt_family="AT-rich"
7678..7792
/rpt_family="ERV1"
7878..8058


```

repeat_region      /rpt_family="ERY1"
                    8162..8483
repeat_region      /rpt_family="ERY1"
                    8491..8791
repeat_region      /rpt_family="Alu"
                    8937..9063
repeat_region      /rpt_family="Alu"
                    9067..9669
repeat_region      /rpt_family="ERV1"
                    9682..9625
repeat_region      /rpt_family="ERV1"
                    9871..10540
repeat_region      /rpt_family="L1"
                    10546..10733
repeat_region      /rpt_family="L1"
                    10734..11190
repeat_region      /rpt_family="L1"
                    15973..16198
repeat_region      /rpt_family="ERV1"
                    17703..17734
repeat_region      /rpt_family="(TG)n"
                    1777..18013
repeat_region      /rpt_family="L2"
                    18844..18871
repeat_region      /rpt_family="AT_rich"
                    18989..19088
repeat_region      /rpt_family="(TA)n"
                    19113..19140
repeat_region      /rpt_family="(TATATG)n"
                    19350..19537
repeat_region      /rpt_family="MALR"
                    19553..19710
repeat_region      /rpt_family="L1"
                    19748..19770
repeat_region      /rpt_family="AT_rich"
                    19940..20478
repeat_region      /rpt_family="ERV1"
                    20453..20495
repeat_region      /rpt_family="(TCTA)n"
                    20533..20930
repeat_region      /rpt_family="L1"
                    20643..20688
repeat_region      /rpt_family="AT_rich"
                    20938..21055
repeat_region      /rpt_family="L1"
                    22130..22220
repeat_region      /rpt_family="L2"
                    22440..22559
repeat_region      /rpt_family="L1"
                    22677..23841
repeat_region      /rpt_family="MER2_type"
                    23210..23257
repeat_region      /rpt_family="AT_rich"
                    23351..23432
repeat_region      /rpt_family="AT_rich"
                    23846..24390
repeat_region      /rpt_family="ERV1"
                    24466..24606
repeat_region      /rpt_family="L1"
                    24607..24756
repeat_region      /rpt_family="Alu"
                    24758..25541
repeat_region      /rpt_family="ERV1"
                    25600..26211
repeat_region      /rpt_family="L1"
                    26212..26607
repeat_region      /rpt_family="ERV1"
                    26608..27012
Query Match      95.28; Score 94.2; DB 9; Length 169286;
Best Local Similarity 97.08; Pred. No. 5,4e-22;
Matches 96; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```

```

QY      1 TCACCCGAGCGCAAGCAATGTGGCATGCTGTGAAGACCACTGAATCCAGCA 60
      |||
Db      83577 TCACCCCTGAGCAAGCAAGCATGTGTGTAAGACCACTGAATCCAGCA 83518
      |||
QY      61 GCCCAGACCCCTTTCTTGTGTGTCAGAAAGCGCGGAAA 99
      |||
Db      83517 GCCCAGACCCCTTTCTTGTGTGTCAGAAAGCGCGGAAA 83479
      |||

RESULT 11
AC024033/c 169462 bp DNA 1linear HTG 24 JAN-2002
LOCUS Homo sapiens chromosome 3 clone RP11-215L17, WORKING DRAFT
DEFINITION
SEQUENCE, 18 unordered pieces.
AC024033
AC024033.5 GI:18308831
VERSION HTG; HTGS_PHASE1; HTGS_DRAFT.
KEYWORDS Homo sapiens.
SOURCE Homo sapiens.
ORGANISM Homo sapiens; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
REFERENCE
AUTHORS Waterston, R.H.
TITLE The sequence of Homo sapiens clone
JOURNAL Unpublished
AUTHORS 2 (bases 1 to 169462)
TITLE Direct Submission
AUTHORS Waterston, R.H.
JOURNAL Submitted (20-FEB-2000) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
On Jan 24, 2002 this sequence version replaced gi:8572509.

COMMENT
----- Genome Center -----
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: http://genome.wustl.edu/gsc/index.shtml
Contact: submissions@wustl.edu
Project Information
Center project name: H_NH0215L17
----- Summary Statistics -----
Sequencing vector: M13; 89%
Sequencing vector: plasmid; 11%
Chemistry: Dye-primer ET; 89% of reads
Chemistry: Dye-terminator Big Dye; 11% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 162787 bases at least Q40
Consensus quality: 165310 bases at least Q30
Consensus quality: 166295 bases at least Q20
Insert size: 146000; agarose-fp
Insert size: 169084; sum-of-contigs
Quality coverage: 4.10 in Q20 bases; agarose-fp
Quality coverage: 4.67 in Q20 bases; sum-of-contigs
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 18 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
* 1 1088: contig of 1088 bp in length
* 1089 1188: gap of unknown length
* 1189 2342: contig of 1154 bp in length
* 2343 2442: gap of unknown length
* 2443 4017: contig of 1575 bp in length
* 4018 4117: gap of unknown length
* 4118 5126: contig of 1009 bp in length
* 5127 5227: gap of unknown length
* 5227 8869: contig of 3643 bp in length
* 8869 8870: gap of unknown length

```

```

*      8970      13746: contig of 4777 bp in length
*      13747      13846: gap of unknown length
*      13847      18317: contig of 4471 bp in length
*      18318      18417: gap of unknown length
*      18418      22593: contig of 4176 bp in length
*      22594      22694: gap of unknown length
*      22694      27521: contig of 4828 bp in length
*      27522      27621: gap of unknown length
*      27622      31811: contig of 4190 bp in length
*      31812      31911: gap of unknown length
*      31912      41611: contig of 9700 bp in length
*      41612      41711: gap of unknown length
*      41712      55897: contig of 14186 bp in length
*      55898      55998: gap of unknown length
*      55998      71981: contig of 15983 bp in length
*      71981      72081: gap of unknown length
*      72081      90541: contig of 18461 bp in length
*      90542      90641: gap of unknown length
*      90642      107717: contig of 17076 bp in length
*      107718      107817: gap of unknown length
*      107818      124486: contig of 16669 bp in length
*      124487      124587: gap of unknown length
*      124587      144116: contig of 19530 bp in length
*      144117      144216: gap of unknown length
*      144217      169462: contig of 25246 bp in length.

```

```

FEATURES
source
1..169462
Location/Qualifiers

```

```

misc_feature /note="assembly_name:Contig5"
1189..2342
misc_feature /note="assembly_name:Contig3"
2443..4017
misc_feature /note="assembly_name:Contig14"
4118..5126
misc_feature /note="assembly_name:Contig15
clone_end:SP6
vector_side:left"
5227..8869
misc_feature /note="assembly_name:Contig16"
8970..13746
misc_feature /note="assembly_name:Contig17"
13847..18317
misc_feature /note="assembly_name:Contig18"
18418..22593
misc_feature /note="assembly_name:Contig19"
22694..27521
misc_feature /note="assembly_name:Contig20"
27622..31811
misc_feature /note="assembly_name:Contig21"
31912..41611
misc_feature /note="assembly_name:Contig22"
41712..55897
misc_feature /note="assembly_name:Contig23"
55998..71981
misc_feature /note="assembly_name:Contig24
clone_end:T7
vector_side:right"
72081..90541
misc_feature /note="assembly_name:Contig25"
90642..107717
misc_feature /note="assembly_name:Contig26"
107818..124486
misc_feature /note="assembly_name:Contig27"
124587..144116
misc_feature /note="assembly_name:Contig28"
144217..169462
misc_feature /note="assembly_name:Contig29"
169462..25246
BASE COUNT 55972 a 30198 c 29904 g 51687 t 1701 others
ORIGIN

```

```

Query Match          95.2%; Score 94.2; DB 2; Length 169462;
Best Local Similarity 97.0%; Pred. No. 5,4e-22;
Matches 96; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```

```

OY      1 TCACCCCTGAGCGCAAGCAATGTTGGCAGTCGTGTAAGGACACCTAGATCAGCA 60
          |||
DB      59130 TCACCCCTGAGCGCAAGCAATGTTGGCAGTCGTGTAAGGACACCTAGATCAGCA 59071
          |||
OY      61 GCCCAGACCCCTTTGTTGGTCAAGAAAGCGGGGAA 99
          |||
DB      59070 GCCCAGACCCCTTTGTTGGTCAAGAAAGCGGGGAA 59032
          |||

```

```

RESULT 12
AC019193
DEFINITION Homo sapiens BAC clone RP11-335L23 from 4, complete sequence.
ACCESSION AC019193
VERSION AC019193.7 GI:15778762
KEYWORDS HTG.
SOURCE Homo sapiens.
ORGANISM Homo sapiens.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 177720)
Sulston,J.E. and Waterston,R.
Toward a complete human genome sequence
Genome Res. 8 (11), 1097-1108 (1998)
98063792
MEDLINE
PUBMED
9847074

```

```

REFERENCE
2 (bases 1 to 177720)
Edwards,J., Scott,K. and Cordum,H.
The sequence of Homo sapiens BAC clone RP11-335L23
Unpublished (2001)
3 (bases 1 to 177720)
Waterston,R.H.
Direct Submission
Submitted (30-DEC-1999) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
4 (bases 1 to 177720)
Waterston,R.H.
Direct Submission
Submitted (26-SEP-2001) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
5 (bases 1 to 177720)
Waterston,R.
Direct Submission
Submitted (09-JAN-2002) Department of Genetics, Washington
University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
On Sep 26, 2001 this sequence version replaced gi.13606040.

```

```

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: http://genome.wustl.edu/gsc
Contact: saplens@watson.wustl.edu
Summary Statistics
Center project name: H_NH0335123

```

NOTICE: This sequence may not represent the entire insert of this clone. It may be shorter because we only sequence overlapping clone sections once, or longer because we provide a small overlap between neighboring data submissions.

This sequence was finished as follows unless otherwise noted: all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by

restriction digest.

MAPPING INFORMATION: Mapping information for this clone was provided by Dr. John D. McPherson, Department of Genetics, Washington University, St. Louis MO. For additional information about the map position of this sequence, see <http://genome.wustl.edu/gsc>

SOURCE INFORMATION:

The RPCT-11 human BAC library was made from the blood of one male donor, as described by Osoegawa, K., Woon, P. Y., Zhao, B., Firengen, E., Tateno, M., Catanesi, J. J. and de Jong, P. J. (1998) An improved approach for construction of bacterial artificial chromosome libraries. Genomics 51:1-8. The clone may be obtained either from Research Genetics, Inc. (<http://www.resgen.com>) or Pieter de Jong and coworkers at the Roswell Park Cancer Institute (<http://pacpac.med.buffalo.edu>)

VECTOR: pBACe3.6

NEIGHBORING SEQUENCE INFORMATION:

The clone sequenced to the left is RP11-188P17. Actual start of this clone is at base position 1 of RP11-335L23; actual end is at base position 177720 of RP11-335L23.

Data from AC080070 and AC079766 was used to finish this clone, AC019193. Polymorphisms have been identified between AC019193 and AC080070.

FEATURES

Source

1. 177720

/organism="Homo sapiens"

/db_xref="taxon:9606"

/chromosome="4"

/map="4"

/clone="RP11-335L23"

/clone_lib="RPCT-11"

1050. 1350

/rpt_family="Alu"

2334. 2353

/rpt_family="(CA)n"

4076. 4120

/rpt_family="Alu"

4121. 5207

/rpt_family="L1"

5085. 5089

/note="similar to Homo sapiens EST B6994784 (NID:91439865)"

5204. 5835

/rpt_family="L1"

6169. 6237

/rpt_family="L1"

6238. 6546

/rpt_family="Alu"

6516. 6550

/rpt_family="A-rich"

6547. 6737

/rpt_family="L1"

6738. 7036

/rpt_family="Alu"

6860. 6881

/rpt_family="AT-rich"

7037. 7112

/rpt_family="L1"

7163. 8060

/rpt_family="L1"

8063. 8444

/rpt_family="MALR"

8506. 8814

/rpt_family="Alu"

8788. 8814

/rpt_family="AT-rich"

8972. 9283

/rpt_family="Alu"

9254. 9283

/rpt_family="(A)n"

repeat_region

9284. 9306

/rpt_family="(TA)n"

repeat_region

9306. 9334

/rpt_family="(CA)n"

repeat_region

9403. 9557

/rpt_family="L1"

repeat_region

9555. 9666

/rpt_family="L1"

repeat_region

9672. 10122

/rpt_family="L1"

repeat_region

9831. 9902

/rpt_family="GA-rich"

repeat_region

10150. 10275

/rpt_family="L1"

repeat_region

10251. 10433

/rpt_family="L1"

repeat_region

10288. 10603

/rpt_family="L1"

repeat_region

10608. 10660

/rpt_family="MERL-type"

repeat_region

10859. 11020

/rpt_family="L1"

repeat_region

11153. 11521

/rpt_family="MALR"

misc_feature

12033. 12483

/note="match to EST AA309275 (NID:91961665)"

repeat_region

12881. 12581

/rpt_family="Alu"

repeat_region

12600. 12624

/rpt_family="(T)n"

repeat_region

13032. 13102

/rpt_family="MIR"

repeat_region

13105. 13540

/rpt_family="L1"

repeat_region

13522. 13560

/rpt_family="AT-rich"

repeat_region

14010. 14262

/rpt_family="L1"

repeat_region

14249. 15274

/rpt_family="L1"

repeat_region

15275. 15494

/rpt_family="MERL-type"

repeat_region

15495. 15633

/rpt_family="L1"

repeat_region

15836. 15904

/rpt_family="MER2-type"

repeat_region

15895. 16327

/rpt_family="MER2-type"

repeat_region

16329. 16885

/rpt_family="L1"

repeat_region

16886. 17181

/rpt_family="Alu"

repeat_region

17182. 17455

/rpt_family="L1"

repeat_region

17456. 17763

/rpt_family="Alu"

repeat_region

17741. 17773

/rpt_family="(CAAA)n"

repeat_region

17764. 17827

/rpt_family="L1"

repeat_region

17828. 18120

/rpt_family="Alu"

repeat_region

18121. 18805

/rpt_family="L1"

repeat_region

18792. 20289

/rpt_family="L1"

repeat_region

24661. 24699

/rpt_family="ERV1"

repeat_region

24684. 24981

/rpt_family="ERV1"

Query Match 95.2%; Score 94.2; DB 9; Length 177720;
Best Local Similarity 97.0%; Pred. No. 5.4e-22;

27 1 TCACCCCTGAGCGCAAGGCAATGTTGGCATGCTGGTAAAGGACCACCTAGAATCCAGCA 60

/organism="Homo sapiens"

```
/db_xref="taxon:9606"
/chromosome="21"
/map="21q22.3"
/clone="p160L9, 5' partial"
/clone_1lb="RPC11,3-5 PAC library"
/note="Accession No. AF121897"
73410..243533
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="21"
/map="21q22.3"
/clone="p70124"
/clone_1lb="RPC11,3-5 PAC library"
/note="Accession No. AF064860"
224137..>340000
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="21"
/map="21q22.3"
/clone="p206A10, 3' partial"
/clone_1lb="RPC11,3-5 PAC library"
/note="Accession No. AF121782"
286628..>340000
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="21"
/map="21q22.3"
/clone="BAC-291B3, 3' partial"
/clone_1lb="PAC library"
/note="Accession No. AF064857"
1088..1566
/note="LIM4"
/rpt_family="LINE/L1"
/rpt_type=DISPERSED
2398..2495
/note="(TCTA)n"
/rpt_family="Simple_repeat"
/rpt_type=TANDEM
complement(2496..2878)
/note="TREC"
/rpt_family="LTR/MaLR"
/rpt_type=DISPERSED
2879..2931
/note="(TCTA)n"
/rpt_family="Simple_repeat"
/rpt_type=TANDEM
2932..3004
/note="(TCCA)n"
/rpt_family="Simple_repeat"
/rpt_type=TANDEM
3005..3074
/note="(TCCA)n"
/rpt_family="Simple_repeat"
/rpt_type=TANDEM
3473..19157
/gene="SH3BGR"
join(<3473..3565,12649..12678,15331..15460,18739..19157)
/gene="SH3BGR"
join(3473..3565,12649..12678,15331..15426)
/partial
/gene="SH3BGR"
/note="Accession No. X93498"
/codon_start=1
/product="21-Glutamic Acid Rich protein 21-GARP"
/db_xref="GI:7717370"
/translation="GSEKAEGGETEAQEGSEDVGNLPDAQEKNEEGETAETTEE
IAMEGAEGAEFEETAECEEGEDEDS"
3473..3565
/gene="SH3BGR"
/number=4
5542..5829
/note="AluX"
```

```
/rpt_family="SINE/Alu"
/rpt_type=DISPERSED
complement(6887..7067)
/note="L1MD3"
/rpt_family="LINE/L1"
/rpt_type=DISPERSED
complement(8482..8762)
/note="AluJo"
/rpt_family="SINE/Alu"
/rpt_type=DISPERSED
complement(8865..8964)
/note="L2"
/rpt_family="LINE/L2"
/rpt_type=DISPERSED
9131..9424
/note="AluSc"
/rpt_family="SINE/Alu"
/rpt_type=DISPERSED
complement(9723..9882)
/note="L2"
/rpt_family="LINE/L2"
/rpt_type=DISPERSED
complement(9989..10077)
/note="L2"
/rpt_family="LINE/L2"
/rpt_type=DISPERSED
complement(10078..10429)
/note="TREC"
/rpt_family="LTR/MaLR"
/rpt_type=DISPERSED
complement(10430..10651)
/note="L2"
/rpt_family="LINE/L2"
/rpt_type=DISPERSED
10980..11000
/note="(TTTA)n"
/rpt_family="Simple_repeat"
/rpt_type=TANDEM
complement(11002..11282)
/note="AluSp"
/rpt_family="SINE/Alu"
/rpt_type=DISPERSED
complement(11315..11556)
/note="AluJo"
/rpt_family="SINE/Alu"
/rpt_type=DISPERSED
11702..11753
/note="MIR"
/rpt_family="SINE/MIR"
/rpt_type=DISPERSED
12649..12678
/gene="SH3BGR"
/number=5
complement(13115..13425)
/note="AluX"
/rpt_family="SINE/Alu"
/rpt_type=DISPERSED
complement(13598..13892)
/note="AluY"
/rpt_family="SINE/Alu"
/rpt_type=DISPERSED
14013..14131
/note="(CA)n"
/rpt_family="Simple_repeat"
```

Query Match 95.2%; Score 94.2; DB 9; Length 340000;
Best Local Similarity 97.0%; Pred. No. 5.6e-22;
Matches 96; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 TCACCCCTGAGCGCAAGGCAATGTTGGCATGCTGTAAGACCACTGGAATCCAGCA 60
DB 247147 TCACCTCTGAGCGCAAGGCAATGTTGGCATGCTGTAAGACCACTGGAATCCAGCA 247088

0y 61 GCCCAGACCCCTTCTTGTGTCAGAAAGCGGGA 99
|||||
DB 247087 GCCCGACCCCTTCTTGTGTCAGAAAGCGGGA 247049

RESULT 15
AC093393 120450 bp DNA linear PRI 09-JAN-2002
LOCUS Homo sapiens BAC clone RP11-816B18 from 2, complete sequence.
AC093393
AC093393
AC093393.3 GI:18042521
VERSION
KEYWORDS
SOURCE
ORGANISM
Homo sapiens.
Homo sapiens.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.

REFERENCE
AUTHORS
TITLE
JOURNAL
Toward a complete human genome sequence
Genome Res. 8 (11), 1097-1108 (1998)
MEDLINE
99063792
PUBMED
9847074

REFERENCE
AUTHORS
TITLE
JOURNAL
Unpublished (2001)
REFERENCE
AUTHORS
TITLE
JOURNAL
Submitted (21-AUG-2001) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA

REFERENCE
AUTHORS
TITLE
JOURNAL
Submitted (03-JAN-2002) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA

REFERENCE
AUTHORS
TITLE
JOURNAL
Submitted (09-JAN-2002) Department of Genetics, Washington
University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
On Jan 3, 2002 this sequence version replaced gi:16874923.
----- Genome Center
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: <http://genome.wustl.edu/gsc>
Contact: sapiens@wustl.edu
----- Summary Statistics
Center project name: H_NH0816B18

NOTICE: This sequence may not represent the entire insert of this
clone. It may be shorter because we only sequence overlapping
clone sections once, or longer because we provide a small overlap
between neighboring data submissions.

This sequence was finished as follows unless otherwise noted:
all regions were double stranded, sequenced with an alternate
chemistry, or covered by high quality data (i.e., phred quality >
30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by sequence
from more than one subclone; and the assembly was confirmed by
restriction digest.

MAPPING INFORMATION:
Mapping information for this clone was provided by Dr. John D.
McHersom, Department of Genetics, Washington University, St. Louis
MO. For additional information about the map position of this
sequence, see <http://genome.wustl.edu/gsc>

SOURCE INFORMATION:

FEATURES

SOURCE

The RPCT-11 human BAC library was made from the blood of one male
donor, as described by Osoegawa, K., Moon, P. Y., Zhao, B., Frengen, E.,
Tateno, M., Catanese, J. J., and de Jong, P. J. (1996) An improved
approach for construction of bacterial artificial chromosome
libraries. *Genomics* 51:1-8. The clone may be obtained either from
Research Genetics, Inc. (<http://www.resgen.com>) or Pletter de Jong
and coworkers at the Roswell Park Cancer Institute
(<http://bacpac.med.buffalo.edu>)
VECTOR: pBAC3.6
NEIGHBORING SEQUENCE INFORMATION:
The clone sequenced to the left is RP11-330A12; the clone sequenced
to the right is RP11-332N21, 2000 bp overlap. Actual start of this
clone is at base position 1 of RP11-816B18; actual end is at base
position 64829 of RP11-332N21.

Data from AC017050 was used to finish AC093393. Polymorphisms have
been identified between AC017050, AC069303 and AC093393. There is
single plasmid coverage from 114395 to 114413.

Location/Qualifiers

1..120450
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="2"
/map="2"
/clone="RP11-816B18"
/clone_11b="RPCT-11"
1..79
/rpt_family="Alu"
repeat_region
82..143
/rpt_family="MALR"
repeat_region
929..1108
/rpt_family="L2"
1112..1374
/rpt_family="MALR"
1291..1292
/note="similar to Homo sapiens EST AL537709
(NID:g12801202)"
complement(1344..1345)
/note="similar to Homo sapiens EST BF215435
(NID:g11109021)"
complement(1344..1345)
/note="similar to Homo sapiens EST BG502858
(NID:g13464375)"
complement(1344..1345)
/note="similar to Homo sapiens EST BG506376
(NID:g13467893)"
complement(1344..1345)
/note="similar to Homo sapiens EST BG532982
(NID:g13524521)"
1377..2068
/note="similar to Homo sapiens EST A0117318
(NID:g10932276)"
1384..2066
/note="similar to Homo sapiens EST B1549051
(NID:g15436363)"
1385..1693
/note="similar to Homo sapiens EST BG506376
(NID:g13467893)"
1385..1691
/note="similar to Homo sapiens EST BF215435
(NID:g11109021)"
1385..1690
/note="similar to Homo sapiens EST BG532982
(NID:g13524521)"
1385..1480
/note="similar to Homo sapiens EST BG502858
(NID:g13464375)"
1395..2350
/note="similar to EST BM047694 (NID:g16776961)"
1395..1767
/note="similar to EST BM044716 (NID:g16773983)"
1397..2113
/note="similar to Homo sapiens EST BF786425

Search completed: May 2, 2003, 12:52:49
Job time : 826.993 secs

```
misc_feature (NTD:g10207623)"
1398..1668 /note="similar to Homo sapiens EST BF678628
(NID:g11952523)"
misc_feature 1401..2046 /note="similar to Homo sapiens EST BG257033
(NID:g12766849)"
misc_feature 1401..1576 /note="similar to Homo sapiens EST AW954663
(NID:g8144346)"
misc_feature 1403..1867 /note="similar to Homo sapiens EST AA309125
(NID:g1961450)"
misc_feature 1404..1876 /note="similar to Homo sapiens EST BE960674
(NID:g10571483)"
misc_feature 1404..1693 /note="similar to Homo sapiens EST BF028041
(NID:g10735662)"
misc_feature 1405..2220 /note="similar to Homo sapiens EST BG677927
(NID:g13909324)"
misc_feature 1405..1974 /note="similar to Homo sapiens EST BG292090
(NID:g13050544)"
misc_feature 1410..2273 /note="similar to Homo sapiens EST AL537709
(NID:g12801202)"
misc_feature 1410..1794 /note="similar to Homo sapiens EST BG006478
(NID:g12449699)"
misc_feature 1414..2462 /note="similar to Homo sapiens EST AL513628
(NID:g12777122)"
misc_feature 1416..2116 /note="similar to Homo sapiens EST EM012067
(NID:g16526421)"
misc_feature 1420..2125 /note="similar to Homo sapiens EST AL515803
(NID:g12779296)"
misc_feature 1437..1915 /note="similar to Homo sapiens EST BI258319
(NID:g14814547)"
misc_feature 1437..1911 /note="similar to Homo sapiens EST BE295078
(NID:g9178623)"
misc_feature 1437..1759 /note="similar to Homo sapiens EST BI114149
(NID:g14565050)"
misc_feature 1437..1661 /note="similar to Homo sapiens EST AA300977
(NID:g1953330)"
misc_feature 1443..2034 /note="similar to Homo sapiens EST BG035997
(NID:g12430691)"
misc_feature 1446..2158 /note="similar to Homo sapiens EST BG618127
(NID:g13669498)"
misc_feature 1510..2042 /note="similar to Homo sapiens EST BG502858
(NID:g13464375)"
```

Query Match 94.1%; Score 93.2; DB 9; Length 120450;
Best Local Similarity 96.9%; Pred. No. 1.2e-21;
Matches 95; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```
QY 1 TCACCCCTGAGCGCAAGCAATGTTGGCATGCTGTAAGGACCACTAGAAATCCAGCA 60
DB 24603 TCACCCCTGAGCTCAAGGCAATGTTGGCATGCTGTAAGGACCACTAGAAATCCAGCA 24662
QY 61 GCCCAGACCCCTTCTTTGTGTCAGAAAGGGGGA 98
DB 24663 GCCCAGACCCCTTCTTTGTGTCAGAAAGGGGGA 24700
```


GenCore version 5.1.5
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 2, 2003, 11:21:20 ; Search time 157.717 Seconds
(without alignments)
1413.592 Million cell updates/sec

Title: US-09-719-554-3_COPY_2502_2600

Perfect score: 99
Sequence: 1 tcacccttgagcgcaaaagc.....tggtcaagaagcgcggaagaa 99

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 112599159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

N.GeneSeq_101002:*

1: /SID2/gcgdata/geneSeq/geneSeqn-emb1/NA1980.DAT:*

2: /SID2/gcgdata/geneSeq/geneSeqn-emb1/NA1981.DAT:*

3: /SID2/gcgdata/geneSeq/geneSeqn-emb1/NA1982.DAT:*

4: /SID2/gcgdata/geneSeq/geneSeqn-emb1/NA1983.DAT:*

5: /SID2/gcgdata/geneSeq/geneSeqn-emb1/NA1984.DAT:*

6: /SID2/gcgdata/geneSeq/geneSeqn-emb1/NA1985.DAT:*

7: /SID2/gcgdata/geneSeq/geneSeqn-emb1/NA1986.DAT:*

8: /SID2/gcgdata/geneSeq/geneSeqn-emb1/NA1987.DAT:*

9: /SID2/gcgdata/geneSeq/geneSeqn-emb1/NA1988.DAT:*

10: /SID2/gcgdata/geneSeq/geneSeqn-emb1/NA1989.DAT:*

11: /SID2/gcgdata/geneSeq/geneSeqn-emb1/NA1990.DAT:*

12: /SID2/gcgdata/geneSeq/geneSeqn-emb1/NA1991.DAT:*

13: /SID2/gcgdata/geneSeq/geneSeqn-emb1/NA1992.DAT:*

14: /SID2/gcgdata/geneSeq/geneSeqn-emb1/NA1993.DAT:*

15: /SID2/gcgdata/geneSeq/geneSeqn-emb1/NA1994.DAT:*

16: /SID2/gcgdata/geneSeq/geneSeqn-emb1/NA1995.DAT:*

17: /SID2/gcgdata/geneSeq/geneSeqn-emb1/NA1996.DAT:*

18: /SID2/gcgdata/geneSeq/geneSeqn-emb1/NA1997.DAT:*

19: /SID2/gcgdata/geneSeq/geneSeqn-emb1/NA1998.DAT:*

20: /SID2/gcgdata/geneSeq/geneSeqn-emb1/NA1999.DAT:*

21: /SID2/gcgdata/geneSeq/geneSeqn-emb1/NA2000.DAT:*

22: /SID2/gcgdata/geneSeq/geneSeqn-emb1/NA2001A.DAT:*

23: /SID2/gcgdata/geneSeq/geneSeqn-emb1/NA2001B.DAT:*

24: /SID2/gcgdata/geneSeq/geneSeqn-emb1/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	99	100.0	575	23	AA572226
2	99	100.0	2049	23	AA590991
3	99	100.0	4805	23	AA573604
4	99	100.0	4805	23	AA573910
5	99	100.0	4805	23	AA574652
6	99	100.0	4805	23	AA584191
7	99	100.0	6507	23	AA573607
8	99	100.0	6507	23	AA573914
9	99	100.0	6507	23	AA592655

10	99	100.0	8279	23	AA576474	DNA encoding novel
11	99	100.0	8294	21	AA584209	DNA encoding novel
12	99	100.0	10499	21	AA597929	Human retroviral S
13	99	100.0	56093	24	ABL61744	Human adenocarcino
14	97.4	98.4	1687	23	AA572234	DNA encoding novel
15	97.4	98.4	1729	23	AA577018	DNA encoding novel
16	94.2	95.2	1493	23	AA567606	DNA encoding novel
17	94.2	95.2	3831	23	AA571727	DNA encoding novel
18	92.6	93.5	865	23	AA572232	DNA encoding novel
19	92.6	93.5	1307	23	AA571726	DNA encoding novel
20	92.6	93.5	46340	21	ABN97978	Human retroviral s
21	91	91.9	1799	23	AA577017	DNA encoding novel
22	91	91.9	23855	22	AA199411	Human excretory re
23	91	91.9	23855	22	AA163761	Human kidney relat
24	89.4	90.3	510	23	AA576203	DNA encoding novel
25	89.4	90.3	510	23	AA577307	DNA encoding novel
26	89.4	90.3	510	23	AA591839	DNA encoding novel
27	89.4	90.3	1081	23	AA577016	DNA encoding novel
28	87.8	88.7	309	23	AA592668	DNA encoding novel
29	87.8	88.7	435	23	AA592494	DNA encoding novel
30	87.8	88.7	435	23	AA592664	DNA encoding novel
31	87.8	88.7	852	23	AA572224	DNA encoding novel
32	87.8	88.7	954	23	AA572223	DNA encoding novel
33	87.8	88.7	954	23	AA590990	DNA encoding novel
34	87.8	88.7	1539	23	AA573691	DNA encoding novel
35	87.8	88.7	3205	23	AA572236	DNA encoding novel
36	86.2	87.1	1719	23	AA581560	DNA encoding novel
37	85.8	86.7	924	23	AA572224	DNA encoding novel
38	85.8	86.7	1846	23	AA573605	DNA encoding novel
39	85.8	86.7	1846	23	AA584192	DNA encoding novel
40	85.2	86.1	804	23	AA572229	DNA encoding novel
41	85.2	86.1	867	23	AA576195	DNA encoding novel
42	85.2	86.1	1249	23	AA573698	DNA encoding novel
43	75.8	76.6	1518	23	AA564941	DNA encoding novel
44	75.8	76.6	3331	22	AA522442	Human cDNA encodin
45	75.8	76.6	3690	22	AA522678	Human cDNA encodin

ALIGNMENTS

RESULT 1	AA572226	standard; cDNA; 575 BP.
ID	AA572226	
AC	AA572226	
DT	13-FEB-2002	(first entry)
DE	DNA encoding novel human diagnostic protein #8030.	
KW	Human; chromosome mapping; gene mapping; gene therapy; forensic;	
KW	food supplement; medical imaging; diagnostic; genetic disorder; ss.	
XX	Homo sapiens.	
OS		
PN	WO200175067-A2.	
PD	11-OCT-2001.	
PF	30-MAR-2001; 2001MO-US08631.	
PR	31-MAR-2000; 2000US-0540217.	
PR	23-AUG-2000; 2000US-0649167.	
XX	(HYSE-) HYSEQ INC.	
PI	Dymanac RT, Liu C, Tang YT;	
XX		
XX	WPI; 2001-639362/73.	
DR	P-PSDB; ABG08039.	
XX	New isolated polynucleotide and encoded polypeptides, useful in	

PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity -
XX
PS Claim 1; SEQ ID No 8030; 103pp; English.
XX
CC The invention relates to isolated polynucleotide (I) and
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. AAS64197-AAS94564 represent novel human
CC diagnostic coding sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 575 BP; 183 A; 150 C; 140 G; 102 T; 0 other;
XX
Query Match 100.0%; Score 99; DB 23; Length 575;
Best Local Similarity 100.0%; Pred. No. 3.2e-26;
Matches 99; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
OY 1 TCACCCCTGAGCGCAAGCAATGTGGCATGCTGTTAAAGCACCCTAGATCCAGCA 60
DB 23 TCACCCCTGAGCGCAAGCAATGTGGCATGCTGTTAAAGCACCCTAGATCCAGCA 82
OY 61 GCCCAGACCCCTTCTTTGTGTCAGAGAAAGCGCGGAAA 99
DB 83 GCCCAGACCCCTTCTTTGTGTCAGAGAAAGCGCGGAAA 121
XX
RESULT 2
AAS90991
ID AAS90991 standard; cDNA; 2049 BP.
XX
AC AAS90991;
XX
DT 13-FEB-2002 (first entry)
XX
DE DNA encoding novel human diagnostic protein #26795.
XX
KW Human: chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX
OS Homo sapiens.
XX
WO200175067-A2.
XX
PD 11-OCT-2001.
XX
PF 30-MAR-2001; 2001WO-US08631.
XX
PR 31-MAR-2000; 2000US-0540217.
PR 23-AUG-2000; 2000US-0649167.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Drmanac RT, Liu C, Tang YF;
XX
DR WPI: 2001-639362/73.
DR P-PSDB: ABG26804.

XX
PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity -
XX
PS Claim 1; SEQ ID No 26795; 103pp; English.
XX
CC The invention relates to isolated polynucleotide (I) and
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. AAS64197-AAS94564 represent novel human
CC diagnostic coding sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 2049 BP; 612 A; 532 C; 444 G; 461 T; 0 other;
XX
Query Match 100.0%; Score 99; DB 23; Length 2049;
Best Local Similarity 100.0%; Pred. No. 5.1e-26;
Matches 99; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
OY 1 TCACCCCTGAGCGCAAGCAATGTGGCATGCTGTTAAAGCACCCTAGATCCAGCA 60
DB 23 TCACCCCTGAGCGCAAGCAATGTGGCATGCTGTTAAAGCACCCTAGATCCAGCA 82
OY 61 GCCCAGACCCCTTCTTTGTGTCAGAGAAAGCGCGGAAA 99
DB 83 GCCCAGACCCCTTCTTTGTGTCAGAGAAAGCGCGGAAA 121
XX
RESULT 3
AAS73604
ID AAS73604 standard; cDNA; 4805 BP.
XX
AC AAS73604;
XX
DT 13-FEB-2002 (first entry)
XX
DE DNA encoding novel human diagnostic protein #9408.
XX
KW Human: chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX
OS Homo sapiens.
XX
WO200175067-A2.
XX
PD 11-OCT-2001.
XX
PF 30-MAR-2001; 2001WO-US08631.
XX
PR 31-MAR-2000; 2000US-0540217.
PR 23-AUG-2000; 2000US-0649167.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Drmanac RT, Liu C, Tang YF;
XX

DR WPI: 2001-639362/73.
DR P-PSDB: ABG09417.

PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity -

PS Claim 1; SEQ ID No 9408; 103pp; English.

XX The invention relates to isolated polynucleotide (I) and
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. AAS64197-AAS94564 represent novel human
CC diagnostic coding sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.

SQ Sequence 4805 BP; 1400 A; 1173 C; 1009 G; 1223 T; 0 other;

Query Match 100.0%; Score 99; DB 23; Length 4805;
Best Local Similarity 100.0%; Pred. No. 6.9e-26;

Matches 99; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 TCACCCCTGAGCGCAAGCAATGTTGGCATGCTGTAAAGACCAATCCAGCA 60
|||||
DB 2950 TCACCCCTGAGCGCAAGCAATGTTGGCATGCTGTAAAGACCAATCCAGCA 3009

OY 61 GCCCAGACCCCTTTCTTTGTGCTCAAGAAAGCGGGAAA 99
|||||
DB 3010 GCCCAGACCCCTTTCTTTGTGCTCAAGAAAGCGGGAAA 3048

RESULT 4
AAS73910
ID AAS73910 standard; cDNA; 4805 BP.

XX AAS73910;

DT 13-FEB-2002 (first entry)

XX DNA encoding novel human diagnostic protein #9714.

XX Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder; ss.

OS Homo sapiens.

XX WO200175067-A2.

XX 11-OCT-2001.

PD 30-MAR-2001; 2001WO-US08631.

PF 31-MAR-2000; 2000US-0540217.

PR 23-AUG-2000; 2000US-0649167.

XX (HYSE-) HYSEQ INC.
XX

PI Drmanac RT, Liu C, Tang YT;
XX
DR WPI: 2001-639362/73.
DR P-PSDB: ABG09723.

PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity -

PS Claim 1; SEQ ID No 9714; 103pp; English.

XX The invention relates to isolated polynucleotide (I) and
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. AAS64197-AAS94564 represent novel human
CC diagnostic coding sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.

SQ Sequence 4805 BP; 1400 A; 1173 C; 1009 G; 1223 T; 0 other;

Query Match 100.0%; Score 99; DB 23; Length 4805;
Best Local Similarity 100.0%; Pred. No. 6.9e-26;

Matches 99; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 TCACCCCTGAGCGCAAGCAATGTTGGCATGCTGTAAAGACCAATCCAGCA 60
|||||
DB 2950 TCACCCCTGAGCGCAAGCAATGTTGGCATGCTGTAAAGACCAATCCAGCA 3009

OY 61 GCCCAGACCCCTTTCTTTGTGCTCAAGAAAGCGGGAAA 99
|||||
DB 3010 GCCCAGACCCCTTTCTTTGTGCTCAAGAAAGCGGGAAA 3048

RESULT 5
AAS76462
ID AAS76462 standard; cDNA; 4805 BP.

XX AAS76462;

DT 13-FEB-2002 (first entry)

XX DNA encoding novel human diagnostic protein #12266.

XX Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder; ss.

OS Homo sapiens.

XX WO200175067-A2.

XX 11-OCT-2001.

PD 30-MAR-2001; 2001WO-US08631.

PF 31-MAR-2000; 2000US-0540217.

PR 23-AUG-2000; 2000US-0649167.
XX

PA (HYSE-) HYSEQ INC.
XX
PI Drmanac RT, Liu C, Tang YT;
XX
XX WPI; 2001-639362/73.
DR P-PSDB; ABG12275.
XX
XX New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity
XX
XX Claim 1; SEQ ID No 12266; 103bp; English.
XX
XX The invention relates to isolated polynucleotide (I) and
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. AAS64197-AAS94564 represent novel human
CC diagnostic coding sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 4805 BP; 1400 A; 1173 C; 1009 G; 1223 T; 0 other;
Query Match 100.0%; Score 99; DB 23; Length 4805;
Best Local Similarity 100.0%; Pred. No. 6.9e-26;
Matches 99; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TCACCCCTGAGCGCAAGGCAATGTGGCATGCTGTAAAGCACCTGAAATCCAGCA 60
|||||
DB 2950 TCACCCCTGAGCGCAAGGCAATGTGGCATGCTGTAAAGCACCTGAAATCCAGCA 3009
QY 61 GCCCAGACCCCTTCTTGTGTGTCAGAAAGCGGGAAA 99
|||||
DB 3010 GCCCAGACCCCTTCTTGTGTGTCAGAAAGCGGGAAA 3048
RESULT 6
AAS84191
ID AAS84191 standard; cDNA; 4805 BP.
XX
AC AAS84191;
XX
DT 13-FEB-2002 (first entry)
XX
DE DNA encoding novel human diagnostic protein #19955.
XX
KW Human: chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX
OS Homo sapiens.
XX
PN WO200175067-A2.
XX
PD 11-OCT-2001.
XX
PF 30-MAR-2001; 2001MO-US08631.
XX
PR 31-MAR-2000; 2000US-0540217.

PR 23-AUG-2000; 2000US-0649167.
XX
XX (HYSE-) HYSEQ INC.
PA
XX
PI Drmanac RT, Liu C, Tang YT;
XX
XX WPI; 2001-639362/73.
DR P-PSDB; ABG20004.
XX
XX New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity
XX
XX Claim 1; SEQ ID No 19955; 103bp; English.
XX
XX The invention relates to isolated polynucleotide (I) and
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. AAS64197-AAS94564 represent novel human
CC diagnostic coding sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 4805 BP; 1400 A; 1173 C; 1009 G; 1223 T; 0 other;
Query Match 100.0%; Score 99; DB 23; Length 4805;
Best Local Similarity 100.0%; Pred. No. 6.9e-26;
Matches 99; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TCACCCCTGAGCGCAAGGCAATGTGGCATGCTGTAAAGCACCTGAAATCCAGCA 60
|||||
DB 2950 TCACCCCTGAGCGCAAGGCAATGTGGCATGCTGTAAAGCACCTGAAATCCAGCA 3009
QY 61 GCCCAGACCCCTTCTTGTGTGTCAGAAAGCGGGAAA 99
|||||
DB 3010 GCCCAGACCCCTTCTTGTGTGTCAGAAAGCGGGAAA 3048
RESULT 7
AAS73607
ID AAS73607 standard; cDNA; 6507 BP.
XX
AC AAS73607;
XX
DT 13-FEB-2002 (first entry)
XX
DE DNA encoding novel human diagnostic protein #9411.
XX
KW Human: chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX
OS Homo sapiens.
XX
PN WO200175067-A2.
XX
PD 11-OCT-2001.
XX
PF 30-MAR-2001; 2001MO-US08631.

XX 31-MAR-2000; 2000US-0540217.
PR 23-AUG-2000; 2000US-0649167.
XX (HYSE-) HXSEQ INC.
XX Drmanac RT, Liu C, Tang YT;
XX WPI: 2001-639362/73.
DR P-PSDB: ABG09420.
XX
XX New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity -
XX
XX Claim 1; SEQ ID No 9411; 103pp; English.
XX
XX The invention relates to isolated polynucleotide (I) and
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. AAS64197-AAS94564 represent novel human
CC diagnostic coding sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
XX Sequence 6507 BP; 1936 A; 1524 C; 1395 G; 1652 T; 0 other;
SQ
Query Match 100.0%; Score 99; DB 23; Length 6507;
Best Local Similarity 100.0%; Pred. No. 7.6e-26;
Matches 99; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 TCACCCCTGAGCGCAAGCAATGTTGGCATGCTGTAAGAGACCTAGATCCAGCA 60
|||||
DB 4652 TCACCCCTGAGCGCAAGCAATGTTGGCATGCTGTAAGAGACCTAGATCCAGCA 4711
OY 61 GCCCAGACCCCTTCTTGTGTGTCAGAAAGCGGGGAAA 99
|||||
DB 4712 GCCCAGACCCCTTCTTGTGTGTCAGAAAGCGGGGAAA 4750
|||||
RESULT 8
AAS73914
ID AAS73914 standard; cDNA: 6507 BP.
XX
XX AAS73914;
AC
XX
XX 13-FEB-2002 (first entry)
DT
XX
XX DNA encoding novel human diagnostic protein #9718.
DE
XX
XX Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX
XX Homo sapiens.
OS
XX WO200175067-A2.
PN
XX
PD 11-OCT-2001.

XX 30-MAR-2001; 2001WO-US08631.
PF
XX 31-MAR-2000; 2000US-0540217.
PR 23-AUG-2000; 2000US-0649167.
XX (HYSE-) HXSEQ INC.
XX Drmanac RT, Liu C, Tang YT;
XX WPI: 2001-639362/73.
DR P-PSDB: ABG09727.
XX
XX New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity -
XX
XX Claim 1; SEQ ID No 9718; 103pp; English.
XX
XX The invention relates to isolated polynucleotide (I) and
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. AAS64197-AAS94564 represent novel human
CC diagnostic coding sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
XX Sequence 6507 BP; 1936 A; 1524 C; 1395 G; 1652 T; 0 other;
SQ
Query Match 100.0%; Score 99; DB 23; Length 6507;
Best Local Similarity 100.0%; Pred. No. 7.6e-26;
Matches 99; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 TCACCCCTGAGCGCAAGCAATGTTGGCATGCTGTAAGAGACCTAGATCCAGCA 60
|||||
DB 4652 TCACCCCTGAGCGCAAGCAATGTTGGCATGCTGTAAGAGACCTAGATCCAGCA 4711
OY 61 GCCCAGACCCCTTCTTGTGTGTCAGAAAGCGGGGAAA 99
|||||
DB 4712 GCCCAGACCCCTTCTTGTGTGTCAGAAAGCGGGGAAA 4750
|||||
RESULT 9
AAS92665/c
ID AAS92665 standard; cDNA: 6507 BP.
XX
XX AAS92665;
AC
XX
XX 13-FEB-2002 (first entry)
DT
XX
XX DNA encoding novel human diagnostic protein #28469.
DE
XX
XX Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX
XX Homo sapiens.
OS
XX WO200175067-A2.
PN

XX 11-OCT-2001.
PD 30-MAR-2001; 2001MO-US08631.
XX
XX 31-MAR-2000; 2000US-0540217.
PR 23-AUG-2000; 2000US-0649167.
XX
XX (HYSE-) HYSEQ INC.
XX
PI Dmanac RT, Liu C, Tang YT;
XX
XX WPI; 2001-639362/73.
DR P-PSDB; ABG28478.
XX
PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity.
XX
XX Claim 1; SEQ ID No 28469; 103pp; English.
XX
CC The invention relates to isolated polynucleotide (I) and
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. AAS64197-AAS94564 represent novel human
CC diagnostic coding sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 6507 BP; 1652 A; 1395 C; 1524 G; 1936 T; 0 other;
Query Match 100.0%; Score 99; DB 23; Length 6507;
Best Local Similarity 100.0%; Pred. No. 7.6e-26;
Matches 99; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TCACCCCTGAGCGCAAGGCAATGTTGGCATGCTGTAAGGACCACTGAATCCAGCA 60
|||||
DB 1856 TCACCCCTGAGCGCAAGGCAATGTTGGCATGCTGTAAGGACCACTGAATCCAGCA 1797
|||||
QY 61 GCCCAGACCCCTTTCTTTGTGTCAGAAAGCGGGGAAA 99
|||||
DB 1796 GCCCAGACCCCTTTCTTTGTGTCAGAAAGCGGGGAAA 1758
|||||
RESULT 10
AAS76474
ID AAS76474 standard; cDNA; 8279 BP.
XX
AC AAS76474;
XX
DT 13-FEB-2002 (first entry)
XX
DE DNA encoding novel human diagnostic protein #12278.
XX
KM Human: chromosome mapping; gene mapping; gene therapy; forensic;
KM food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX
OS Homo sapiens.

XX
PN WO200175067-A2.
XX
XX 11-OCT-2001.
PD 30-MAR-2001; 2001MO-US08631.
XX
XX 31-MAR-2000; 2000US-0540217.
PR 23-AUG-2000; 2000US-0649167.
XX
XX (HYSE-) HYSEQ INC.
XX
PI Dmanac RT, Liu C, Tang YT;
XX
XX WPI; 2001-639362/73.
DR P-PSDB; ABG12287.
XX
PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity.
XX
XX Claim 1; SEQ ID No 12278; 103pp; English.
XX
CC The invention relates to isolated polynucleotide (I) and
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. AAS64197-AAS94564 represent novel human
CC diagnostic coding sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 8279 BP; 2474 A; 1992 C; 1770 G; 2043 T; 0 other;
Query Match 100.0%; Score 99; DB 23; Length 8279;
Best Local Similarity 100.0%; Pred. No. 8.3e-26;
Matches 99; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TCACCCCTGAGCGCAAGGCAATGTTGGCATGCTGTAAGGACCACTGAATCCAGCA 60
|||||
DB 4652 TCACCCCTGAGCGCAAGGCAATGTTGGCATGCTGTAAGGACCACTGAATCCAGCA 4711
|||||
QY 61 GCCCAGACCCCTTTCTTTGTGTCAGAAAGCGGGGAAA 99
|||||
DB 4712 GCCCAGACCCCTTTCTTTGTGTCAGAAAGCGGGGAAA 4750
|||||
RESULT 11
AAS84209
ID AAS84209 standard; cDNA; 8294 BP.
XX
AC AAS84209;
XX
DT 13-FEB-2002 (first entry)
XX
DE DNA encoding novel human diagnostic protein #20013.
XX
KM Human: chromosome mapping; gene mapping; gene therapy; forensic;
KM food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX

XX	Homo sapiens.
OS	
XX	
PN	W0200175067-A2.
XX	
PD	11-OCT-2001.
XX	
PF	30-MAR-2001; 2001WO-US08631.
XX	
PR	31-MAR-2000; 2000US-0540217.
PR	23-AUG-2000; 2000US-0649167.
XX	
PA	(HYSE-) HYSEQ INC.
P1	Drmamac RT, Liu C, Tang YF;
XX	
DR	WPI; 2001-639362/73.
DR	P-PsDB; ABG20022.
XX	
PT	New isolated polynucleotide and encoded polypeptides, useful in
PT	diagnostics, forensics, gene mapping, identification of mutations
PT	responsible for genetic disorders or other traits and to assess
XX	biodiversity -
XS	Claim 1; SEQ ID No 20013; 103pp; English.

The invention relates to isolated polynucleotide (I) and polypeptide (II) sequences. (I) is useful as hybridisation probes, polymerase chain reaction (PCR) primers, oligomers, and for chromosome and gene mapping, and in recombinant production of (II). The polynucleotides are also used in diagnostics as expressed sequence tags for identifying expressed genes. (I) is useful in gene therapy techniques to restore normal activity of (II) or to treat disease states involving (II). (II) is useful for generating antibodies against it, detecting or quantitating a polypeptide in tissue, as molecular weight markers and as a food supplement. (II) and its binding partners are useful in medical imaging of sites expressing (II). (I) and (II) are useful for treating disorders involving aberrant protein expression or biological activity. The polypeptide and polynucleotide sequences have applications in diagnostics, forensics, gene mapping, identification of mutations responsible for genetic disorders or other traits to assess biodiversity and to produce other types of data and products dependent on DNA and amino acid sequences. AAs64197-AA69454 represent novel human diagnostic coding sequences of the invention.

Note: The sequence data for this patent did not appear in the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pat_sequences.

SQ Sequence 8294 BP; 2476 A; 1994 C; 1774 G; 2050 T; 0 other;

Query Match	100.0%	Score 99	DB 23	Length 8294
Best Local Similarity	100.0%	Pred. No. 8.3e-26		
Matches 99	Conservative 0	Mismatches 0	Indels 0	Gaps 0

Accession	Sequence	Position
OY	TCACCCCTGAGCGAAAGGCAAGTTGGGCAAGCTGGTAAAGACACTAGATCCACA	60
Db	TCACCCCTGAGCGAAAGGCAAGTTGGGCAAGCTGGTAAAGACACTAGATCCACA	4711
OY	61 GCCCAGACCCCTTCTTGTGTGTCAGAAAGGCGGGAAA	99
Db	6112 GCCCAGACCCCTTCTTGTGTGTCAGAAAGGCGGGAAA	4750

RESULT 12

ABN97929

ID ABN97929 standard; DNA; 10499 BP.

XX
AC ABN97929;

XX	01-AUG-2002	(first entry)
DT		

XX	Human retroviral sequence HERV-7q.
DE	

XX

KW Autoimmune disease; HERV-7q; chromosome 7q; immunotherapy;
KW multiple sclerosis; ds.

XX Human retrovirus.

XX
PN W09967395-A1

XX 29-DEC-1999.
PD

XX 23-JUN-1999; 99WO-FR01513.
PF

XX 23-JUN-1998; 98FR-0007920.
PR

XX (INRM) INSERM INST SANTE & RECH MEDICALE. PA

XX Alliel PM, Perin J, Rieger F;
PI

XX WPI; 2000-160587/14.
DR

PT New nucleic acid sequences of human endogenous retrovirus, HERV-7q
PT used for diagnosis, treatment and prevention of autoimmune and
PT neurological diseases -

XX	Claim 3; Fig 1; 225pp; French.
PS	

The present invention relates to new nucleic acid sequences of human endogenous retroviruses, HERV-7q, which is located on chromosome 7q. Regulatory elements associated with HERV-7q may alter expression of other genes (even remote genes) on the same chromosome, inducing immunological and/or neurological changes (which may be pathological or protective/curative). HERV-7q peptides can be used to improve efficiency of the immune response, e.g. in immunotherapy. HERV-7q peptides and their coding sequences can be used in immunogenic or vaccinating compositions, for protection against autoimmune diseases, particularly multiple sclerosis. The peptides may also be used (by sequence comparison) to detect/identify endogenous retroviruses that are abnormally expressed in cancer, neuropathologies or other autoimmune diseases. The present sequence was used to illustrate the invention.

.XX Sequence 10499 BP; 3048 A; 2676 C; 2280 G; 2495 T; 0 other;
SQ

Query Match	100.0%	Score 99	DB 21	Length 10499
Best Local Similarly	100.0%	Pred. No.	9e-26	
Match 99, Conservative	0	Mismatches	0	Gaps 0

[illegible]

QY 61 GCCCAGACCCCTTCTCTGTGTCACGAAGAAGCGGGAAA 99
|||||
Db 2562 GCCCAGACCCCTTCTCTGTGTCACGAAGAAGCGGGAAA 2600
|||||

RESULT 13
ABL61744

ID	ABL61744	standard; DNA; 56093 BP.
XX		

AC ABL61744;
XX

DT 15-MAY-2002 (first entry)
XX

DE Colon adenocarcinoma related gene sequence SEQ ID NO:81.
XX

KW Human; cancer; colon; breast; ovary; oesophagus; kidney; thyroid
 KW stomach; lung; prostate; pancreas; carcinoma; antitumour; cancer

KM cytostatic; gene therapy; antineoplastic; Wilm's tumour; adenocarcinoma,
KM gene; ds.

XX	Homo sapiens
OS	

XX	
PN	W0200194629-A2

XX

PD 13-DEC-2001.
XX
PF 30-MAY-2001; 2001WO-US10838.
XX
PR 05-JUN-2000; 2000US-209473P.
PR 05-JUN-2000; 2000US-209531P.
PR 18-SEP-2000; 2000US-233133P.
PR 18-SEP-2000; 2000US-233617P.
PR 20-SEP-2000; 2000US-234009P.
PR 20-SEP-2000; 2000US-234034P.
PR 20-SEP-2000; 2000US-234052P.
PR 22-SEP-2000; 2000US-234503P.
PR 22-SEP-2000; 2000US-234567P.
PR 22-SEP-2000; 2000US-234923P.
PR 25-SEP-2000; 2000US-234924P.
PR 25-SEP-2000; 2000US-235077P.
PR 25-SEP-2000; 2000US-235082P.
PR 25-SEP-2000; 2000US-235134P.
PR 25-SEP-2000; 2000US-235280P.
PR 26-SEP-2000; 2000US-235637P.
PR 26-SEP-2000; 2000US-235638P.
PR 27-SEP-2000; 2000US-235711P.
PR 27-SEP-2000; 2000US-235720P.
PR 27-SEP-2000; 2000US-235840P.
PR 27-SEP-2000; 2000US-235863P.
PR 28-SEP-2000; 2000US-236028P.
PR 28-SEP-2000; 2000US-236032P.
PR 28-SEP-2000; 2000US-236033P.
PR 28-SEP-2000; 2000US-236034P.
PR 28-SEP-2000; 2000US-236109P.
PR 28-SEP-2000; 2000US-236111P.
PR 29-SEP-2000; 2000US-236842P.
PR 29-SEP-2000; 2000US-236891P.
PR 02-OCT-2000; 2000US-237172P.
PR 02-OCT-2000; 2000US-237173P.
PR 02-OCT-2000; 2000US-237278P.
PR 02-OCT-2000; 2000US-237294P.
PR 02-OCT-2000; 2000US-237295P.
PR 02-OCT-2000; 2000US-237315P.
PR 03-OCT-2000; 2000US-237425P.
PR 03-OCT-2000; 2000US-237598P.
PR 03-OCT-2000; 2000US-237604P.
PR 03-OCT-2000; 2000US-237606P.
PR 03-OCT-2000; 2000US-237608P.
PR 01-NOV-2000; 2000US-244867P.
PR 01-NOV-2000; 2000US-245084P.
XX
PA (AVAL-) AVALON PHARM.
XX
PI Young PE, Augustus M, Carter KC, Ebner R, Endress G, Horrigan S;
PI Soppet DR, Weaver Z;
XX
XX WPI; 2002-188264/24.
XX
PT Screening for anti-neoplastic agent involves exposing cells to a
PT chemical agent to be tested for anti-neoplastic activity, and
PT determining a change in expression of a gene of a signature gene set -
XX
PS Claim 1; SEQ ID 81; 44pp; English.
XX
XX The present invention describes a method (M1) for screening for an
XX anti-neoplastic agent. The method involves exposing cells to a chemical
XX agent to be tested for anti-neoplastic activity, determining a change in
XX expression of at least one gene (I) of a signature gene set, where (I)
XX comprises a sequence (S) selected from 8447 sequences (given in ABL61664
XX to ABL70110), or is at least 95% identical to (S), where a change in
XX expression is indicative of anti-neoplastic activity. (I) has cytoskeletal
XX activity and can be used in gene therapy. M1 can be used for screening
XX an anti-neoplastic agent, and can be used for producing a product which
XX is the data collected with respect to the anti-neoplastic agent as a
XX result of M1, and the data is sufficient to convey the chemical
XX structure and/or properties of the agent. M1 can be used in the
XX treatment of cancer such as colon, breast, stomach, lung, thyroid.

CC oesophageal, ovarian, kidney, prostate or pancreatic cancer,
CC adenocarcinoma, carcinoma, clear cell cancer, infiltrating ductal cancer,
CC infiltrating lobular cancer, squamous cell carcinoma, neuroendocrine
CC carcinoma, papillary carcinoma and Wilms' tumour.
XX
SO Sequence 56093 BP; 16164 A; 12346 C; 10702 G; 16681 T; 0 other;
Query Match 100.0%; Score 99; DB 24; Length 56093;
Best Local Similarity 100.0%; Pred. No. 1,6e-25;
Matches 99; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 TCACCCCTGAGCGCAAGGCAATGTGGCATGCTGTGAAGCACTAGATCCAGCA 60
DB 30502 TCACCCCTGAGCGCAAGGCAATGTGGCATGCTGTGAAGCACTAGATCCAGCA 30561
OY 61 GCCCAGACCCCTTCTTGTGTCAGAAAGCGGGGAA 99
DB 30562 GCCCAGACCCCTTCTTGTGTCAGAAAGCGGGGAA 30600
RESULT 14
AAS72234
ID AAS72234 standard; cDNA; 1687 BP.
XX
AC AAS72234;
XX
DT 13-FEB-2002 (first entry)
XX
DE DNA encoding novel human diagnostic protein #8038.
XX
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX
OS Homo sapiens.
XX
PN W0200175067-A2.
XX
PD 11-OCT-2001.
XX
PF 30-MAR-2001; 2001WO-US08631.
XX
PR 31-MAR-2000; 2000US-0540217.
PR 23-AUG-2000; 2000US-0649167.
XX
PA (HXSE-) HXSEQ INC.
XX
PI Drmanac RT, Liu C, Tang YT;
XX
XX WPI; 2001-639362/73.
XX
DR P-PSDB; ABC08047.
XX
XX
PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity -
XX
XX Claim 1; SEQ ID No 8038; 103pp; English.
XX
XX The invention relates to isolated polynucleotide (I) and
XX polypeptide (II) sequences. (I) is useful as hybridisation probes,
XX polymerase chain reaction (PCR) primers, oligomers, and for chromosome
XX and gene mapping, and in recombinant production of (II). The
XX polynucleotides are also used in diagnostics as expressed sequence tags
XX for identifying expressed genes. (I) is useful in gene therapy techniques
XX to restore normal activity of (II) or to treat disease states involving
XX (II). (II) is useful for generating antibodies against it, detecting or
XX quantitating a polypeptide in tissue, as molecular weight markers and as
XX a food supplement. (II) and its binding partners are useful in medical
XX imaging of sites expressing (II). (I) and (II) are useful for treating
XX disorders involving aberrant protein expression or biological activity.
XX The polypeptide and polynucleotide sequences have applications in
XX diagnostics, forensics, gene mapping, identification of mutations
XX responsible for genetic disorders or other traits to assess biodiversity

CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. AAS64197-AAS94564 represent novel human
CC diagnostic coding sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 1687 BP; 513 A; 434 C; 357 G; 383 T; 0 other;
Query Match 98.4%; Score 97.4; DB 23; Length 1687;
Best Local Similarity 99.0%; Pred. No. 1.8e-25;
Matches 98; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
OY 1 TCACCCCTGAGCGCAAGCAATGTTGGCATGCTGTAAGAGACCCTAGATCCAGCA 60
DB 143 TCACCCCTGAGCACAAAGCAATGTTGGCATGCTGTAAGAGACCCTAGATCCAGCA 202
OY 61 GCCCAGACCCCTTCTTGTGTGTCAGAAAGCGGGGAAA 99
DB 203 GCCCAGACCCCTTCTTGTGTGTCAGAAAGCGGGGAAA 241
RESULT 15
AAS77018
ID AAS77018 standard; cDNA; 1729 BP.
XX
AC AAS77018;
XX
DT 13-FEB-2002 (first entry)
XX
DE DNA encoding novel human diagnostic protein #12822.
XX
KW Human: Chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX
OS Homo sapiens.
XX
PN WO2001/5067-A2.
XX
PD 11-OCT-2001.
XX
PF 30-MAR-2001; 2001WO-US08631.
XX
PR 31-MAR-2000; 2000US-0540217.
PR 23-AUG-2000; 2000US-0649167.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Drmanac RT, Liu C, Tang YT;
XX
DR WPI: 2001-639362/73.
DR P-PSDB; ABG12831.
XX
XX New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity
PS Claim 1; SEQ ID No 12822; 103pp; English.
XX
XX The invention relates to isolated polynucleotide (I) and
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in

CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. AAS64197-AAS94564 represent novel human
CC diagnostic coding sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 1729 BP; 525 A; 429 C; 362 G; 413 T; 0 other;
Query Match 98.4%; Score 97.4; DB 23; Length 1729;
Best Local Similarity 99.0%; Pred. No. 1.8e-25;
Matches 98; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
OY 1 TCACCCCTGAGCGCAAGCAATGTTGGCATGCTGTAAGAGACCCTAGATCCAGCA 60
DB 143 TCACCCCTGAGCACAAAGCAATGTTGGCATGCTGTAAGAGACCCTAGATCCAGCA 202
OY 61 GCCCAGACCCCTTCTTGTGTGTCAGAAAGCGGGGAAA 99
DB 203 GCCCAGACCCCTTCTTGTGTGTCAGAAAGCGGGGAAA 241

Search completed: May 2, 2003, 11:29:37
Job time : 177.717 secs

GenCore version 5.1.5
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 2, 2003, 11:21:20 ; Search time 1140.21 Seconds

(without alignments)
1406.195 Million cell updates/sec

Title: US-09-719-554-3_COPY_2502_2600

Perfect score: 99

Sequence: 1

Scoring table: IDENTITY_NUC

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

EST:
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estinu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hic:*
9: gb_est1:*
10: gb_est2:*
11: gb_hic:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pln:*
21: em_gss_vit:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_tod:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the total score being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	93.2	94.1	432	17	AQ438690
C 2	93.2	94.1	481	17	AQ412419
C 3	92.6	93.5	373	12	BF883154
C 4	89.4	90.3	620	17	AQ624662
C 5	86.2	87.1	565	9	AL698433
C 6	86.2	87.1	592	17	AQ465425

C 7	83.6	84.4	561	9	AA307753
C 8	82.2	83.0	538	17	AQ122695
C 9	81.4	82.2	441	17	B84492
C 10	81.4	82.2	632	17	AQ014046
C 11	53.6	54.1	624	17	AQ060950
C 12	47.2	47.7	515	17	AQ060956
C 13	34.2	34.5	386	14	T96504
C 14	31.6	31.9	482	14	BO560733
C 15	31.6	31.9	683	14	BM941798
C 16	31.6	31.9	722	13	B1851147
C 17	31.3	31.3	816	17	BH565798
C 18	30.8	31.1	918	14	BO070614
C 19	30.4	30.7	659	17	BH107401
C 20	30.3	30.3	664	14	BT112104
C 21	29.6	29.9	188	9	A1877491
C 22	29.6	29.9	310	14	BO974577
C 23	29.6	29.9	511	14	BO262033
C 24	29.4	29.7	572	17	BH738095
C 25	29.4	29.7	688	17	BH731212
C 26	29.4	29.7	707	17	BH532111
C 27	29.4	29.7	756	17	BH703188
C 28	29.2	29.5	380	12	BG572101
C 29	28.6	28.9	742	13	B1888821
C 30	28.6	28.9	765	13	BC972789
C 31	28.4	28.7	720	12	BE872347
C 32	28.2	28.5	544	17	A2957805
C 33	28	28.3	506	17	AQ148423
C 34	28	28.3	623	17	BH516961
C 35	28	28.3	658	13	B1891305
C 36	27.8	28.1	223	17	BH599181
C 37	27.8	28.1	399	10	AW156589
C 38	27.8	28.1	418	17	AQ045384
C 39	27.8	28.1	487	17	AQ126708
C 40	27.8	28.1	535	17	AQ599844
C 41	27.8	28.1	546	17	BH672523
C 42	27.8	28.1	572	17	AQ504740
C 43	27.8	28.1	575	17	BH478321
C 44	27.8	28.1	577	17	BH248929
C 45	27.8	28.1	581	17	AQ389315

ALIGNMENTS

RESULT 1
LOCUS AQ438690/C
DEFINITION HS.5141.A2.F08.SP6E.RPCT-11 Human Male BAC Library Homo sapiens
ACCESSION AQ438690
VERSION AQ438690.1 GI:4550029
KEYWORDS GSS.
SOURCE human.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
AUTHORS 1 (bases 1 to 432)
Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holman,T., Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and Hood,L.
TITLE Sequence-tagged connectors: A sequence approach to mapping and scanning the human genome
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)
MEDLINE 99380589
COMMENT Contact: Mahairas GG, Wallace JC, Hood L
High Throughput Sequencing Center
University of Washington
401 Queen Anne Avenue North, Seattle, WA 98109, USA
Tel: (206) 616-3618
Fax: (206) 616-3887
Email: jwallace@u.washington.edu
Clones are derived from the human BAC library RPCT-11. For BAC library availability, please contact Pieter de Jong

BASE COUNT 108 a 94 c 96 g 74 t 1 others
 ORIGIN

Query Match . 93.5%; Score 92.6; DB 12; Length 373;
 Best Local Similarity 96.0%; Pred. No. 1.8e-20;
 Matches .95; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 TCACCCCTGAGCGCAAGGCAATGTTGGCATGCTGTAAGACCACTAGATCCAGCA 60
 DB 78 TCACCCCTGAGCGCAAGGCAATGTTGGCATGCTGTAAGACCACTAGATCCAGCA 137
 QY 61 GCCCAGACCCCTTTCTTTGTGCTCAAGAAAGCGGGAAA 99
 DB 138 GCCCGATCCCTTTCTTTGTGCTCAAGAAAGCGGGAAA 176

RESULT 4

LOCUS A0624662/c 620 bp DNA linear GSS 16-JUN-1999
 DEFINITION CITBI-EI-2657A12.TR CITBI-EI Homo sapiens genomic clone 2657A12,
 DNA sequence.

ACCESSION A0624662
 VERSION A0624662.1 GI:5087054
 KEYWORDS GSS.

SOURCE

ORGANISM

Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE

AUTHORS

Zhao,S., Adams,M.D., Nierman,W., Malek,J., Shizuya,H., Simon,M. and
 Venter,J.C.

TITLE

Use of BAC End Sequences from Caltech Libraries for Sequence-Ready
 Map Building

JOURNAL

COMMENT

Unpublished (1997)
 Other GSSs: CITBI-EI-2657A12.TF
 Contact: Shaying Zhao, William Nierman, Mark Adams
 Department of Eukaryotic Genomics
 The Institute for Genomic Research
 9712 Medical Center Dr., Rockville, MD 20850
 Tel: 301 838 0200
 Fax: 301 838 0208
 Email: hbe@tigr.org

Clones are available from Research Genetics (info@resgen.com). BAC
 end search page:
http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html.
 Seq primer: M13 Reverse
 Class: BAC ends.

FEATURES

Location/Qualifiers
 1..620

/organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="2657A12"
 /clone_1ib="CITBI-EI"
 /sex="male"
 /cell_type="sperm"
 /note="Vector: pBeloBAC11; Site_1: EcoRI; Site_2: EcoRI;
 Caltech Human BAC Library D"

BASE COUNT

170 a 135 c 120 g 195 t

ORIGIN

Query Match 90.3%; Score 89.4; DB 17; Length 620;
 Best Local Similarity 93.9%; Pred. No. 2.4e-19;
 Matches 93; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 TCACCCCTGAGCGCAAGGCAATGTTGGCATGCTGTAAGACCACTAGATCCAGCA 60
 DB 320 TCACCCCTGAGCGCAAGGCAATGTTGGCATGCTGTAAGACCACTAGATCCAGCA 261

QY 61 GCCCAGACCCCTTTCTTTGTGCTCAAGAAAGCGGGAAA 99
 DB 260 GCCCATACCCCTTTCTTTGTGCTCAAGAAAGCGGGAAA 222

RESULT 5

AL698433/c 565 bp mRNA linear EST 21-MAR-2002
 LOCUS DKFZP686K02108.r1 666 (synonym: hicc3) Homo sapiens cDNA clone

DEFINITION DKFZP686K02108.5, mRNA sequence.
 AL698433
 ACCESSION AL698433.1 GI:19618973
 VERSION EST.

KEYWORDS

SOURCE

ORGANISM

Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 565)
 Bloeker,H., Boecker,M., Brandt,P., Mewes,W., Well,B. and Wiemann
 S.

TITLE

JOURNAL

COMMENT

AUTHORS

EST (Bloeker,H., Boecker,M., Brandt,P., Mewes,W., Well,B. and
 Wiemann,S.)
 Unpublished (1999)
 Contact: Bloeker H
 MIPS

Am Klopferpitz 18a D-82152 Martinsried, Germany
 This is the 5' sequence of the clone insert
 Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
 Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
 sequenced by GBF (National Research Centre for Biotechnology Ltd.,
 Braunschweig/Germany) within the cDNA sequencing consortium of the
 German Genome Project.
 No sl sequence available.

This clone (DKFZP686K02108) is available at the RZPD in Berlin.
 Please contact the RZPD: Ressourcentzntum, Heubnerweg 6, 14059
 Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.

FEATURES

SOURCE

Location/Qualifiers
 1..565

/organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="DKFZP686K02108"
 /clone_1ib="666 (synonym: hicc3)"
 /tissue_type="human skeletal muscle"
 /dev_stage="adult"
 /lab_host="DH10B"
 /note="Vector: pTR1p1ex2; Site_1: SfiI; Site_2: SfiIb;
 cDNA-collection"

BASE COUNT 142 a 126 c 117 g 180 t

ORIGIN

Query Match 87.1%; Score 86.2; DB 9; Length 565;
 Best Local Similarity 91.9%; Pred. No. 2.7e-18;
 Matches 91; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 1 TCACCCCTGAGCGCAAGGCAATGTTGGCATGCTGTAAGACCACTAGATCCAGCA 60
 DB 516 TCACCCCTGAGCGCAAGGCAATGTTGGCATGCTGTAAGACCACTAGATCCAGCA 457

QY 61 GCCCAGACCCCTTTCTTTGTGCTCAAGAAAGCGGGAAA 99
 DB 456 GCCCAGACCCCTTTCTTTGTGCTCAAGAAAGCGGGAAA 418

RESULT 6

LOCUS A0465425/c 592 bp DNA linear GSS 23-APR-1999
 DEFINITION HS_5107_B2_C02_S66E RPCT-11 Human Male BAC Library Homo sapiens
 genomic clone Plate=683 Col=4 Row=F, DNA sequence.

ACCESSION A0465425
 VERSION A0465425.1 GI:4642520
 KEYWORDS GSS.

SOURCE

ORGANISM

Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 592)
 Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T.,

TITLE
Keller, A., Shaker, R., Furlong, J., Young, J., Zhao, S., Adams, M.D. and Hood, L.
JOURNAL MEDLINE
Sequence-tagged connectors: A sequence approach to mapping and scanning the human genome
Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)
99380589
COMMENT
Contact: Mahairas GG, Wallace JC, Hood L
High Throughput Sequencing Center
University of Washington
401 Queen Anne Avenue North, Seattle, WA 98109, USA
Tel: (206) 616-3618
Fax: (206) 616-3887
Email: jwallace@u.washington.edu
Clones are derived from the human BAC library RPCI-11. For BAC library availability, please contact Pieter de Jong (pieter@jcm.med.buffalo.edu). Clones may be purchased from BACPAC Resources (http://bacpac.med.buffalo.edu/ordering_bac.htm) or from Research Genetics (<http://www.htsc.washington.edu>). BAC end Web Server: <http://www.htsc.washington.edu>
Plate: 683 row: F column: 4
Seq primer: SP6
Class: BAC ends
High quality sequence stop: 592.

FEATURES

source
Location/Qualifiers
1..592

BASE COUNT
136 a 158 c 104 g 160 t 34 others
ORIGIN
Query Match 87.1%; Score 86.2; DB 17; Length 592;
Best Local Similarity 91.9%; Pred. No. 2.7e-18;
Matches 91; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
Db 1 TCACCCCTGAGCGCAAGGCAATGTTGGCATGCTGTAAGAGCACTAGATCCAGCA 60
293 TCACCGGGGACACAAAGCAATGTGGGACGCTGTAAGAGCACTAGATCCAGCA 234
Qy 61 GCCCAGACCCCTTTCTTTGTTGTCAGAGAGCGCGGAAA 99
233 GCCCGACACCTTTCTTTGTTGTCAGAGAGCGCGGAAA 195

RESULT 7
AA307753/c 561 bp mRNA linear EST 18-APR-1997
LOCUS EST178614 Colon carcinoma (HCC) cell line Homo sapiens cDNA 5' end,
DEFINITION mRNA sequence.
ACCESSION AA307753
VERSION AA307753.1 GI:1960081
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 561)
Adams, M.D., Kerlavage, A.R., Fleischmann, R.D., Fuldner, R.A., Bult,
C.J., Lee, N.H., Kirkness, E.F., Weissbrock, K.G., Gocayne, J.D., White,
O., Sutton, G., Blake, J.A., Brandon, R.C., Man-Wei, C., Clayton, R.A.,
Cline, T.R., Cotton, M.D., Earle-Hughes, J., Fine, L.D., Fitzgerald,
L.M., Fitzhugh, W.M., Fritchman, J.L., Geoghegan, N.S., Glodex, A.,
Gnelli, C.L., Hanna, M.C., Hedblom, E., Hinkle, P.S., Jr., Kelley, J.M.,
Kelley, J.C., Liu, L.-I., Marmaro, S.M., Merrick, J.M.,
Moreno-Palancas, R.F., McDonald, L.A., Nguyen, D.T., Pelligrino, S.M.,
Phillips, C.A., Ryder, S.E., Scott, J.L., Sauder, D.M., Shirley, R.,

TITLE
Small, K.V., Spriggs, T.A., Utterback, T.R., Weidman, J.F., Li, Y.,
Bednarik, D.P., Cao, L., Cepeda, M.A., Coleman, T.A., Collins, E.J.,
Dimke, D., Feng, D.-F., Ferlie, A., Fischer, C., Hastings, G.A., He, W.W.,
Hu, J.S., Greene, J.M., Gruber, J., Hudson, P., Kim, A.K., Kozak, D.L.,
Kunsch, C., Hungjun, J., Li, H., Weissner, P.S., Olsen, H., Raymond, L.,
Wel, T.F., Weng, J., Xu, C., Yu, G.L., Ruben, S.M., Dillon, P.J., Fannon,
M.R., Rosen, C.A., Haseltine, W.A., Fields, C., Fraser, C.M. and
Venter, J.C.
JOURNAL MEDLINE
Initial assessment of human gene diversity and expression patterns
based upon 83 million nucleotides of cDNA sequence
Nature 377 (6547 suppl), 3-174 (1995)
96026280
COMMENT
Contact: Kerlavage, AR
Bioinformatics
The Institute for Genomic Research
9712 Medical Center Drive, Rockville, MD 20850 USA
Tel: 3016699056
Fax: 3016699423
Email: alkerlav@tigr.org
For clone availability, additional sequence and expression
information related to this EST, please check the TIGR Human Gene
Index (<http://www.tigr.org/tdb/hgi/hgi.html>)
Seq primer: M13 Reverse.

FEATURES

source
Location/Qualifiers
1..361

BASE COUNT
146 a 124 c 114 g 175 t 2 others
ORIGIN
Query Match 84.4%; Score 83.6; DB 9; Length 361;
Best Local Similarity 89.9%; Pred. No. 2e-17;
Matches 89; Conservative 0; Mismatches 10; Indels 0; Gaps 0;
Db 1 TCACCCCTGAGCGCAAGGCAATGTTGGCATGCTGTAAGAGCACTAGATCCAGCA 60
190 TCACCGGGGACACAAAGCAATGTGGGACGCTGTAAGAGCACTAGATCCAGCA 131
Qy 61 GCCCAGACCCCTTTCTTTGTTGTCAGAGAGCGCGGAAA 99
130 GCCCGACACCTTTCTTTGTTGTCAGAGAGCGCGGAAA 92

RESULT 8
A0122695/c 538 bp DNA linear GSS 22-SEP-1998
LOCUS HS_3080_A2_D04_MR_CIT Approved Human Genomic Sperm Library D Homo
DEFINITION sapiens genomic clone Plate=3080 Col=8 Row=G, DNA sequence.
ACCESSION A0122695
VERSION A0122695.1 GI:3499861
KEYWORDS GSS.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 538)
Mahairas, G.G., Wallace, J.C., Smith, K., Swartzell, S., Holzman, T.,
Keller, A., Shaker, R., Furlong, J., Young, J., Zhao, S., Adams, M.D. and
Hood, L.
Sequence-tagged connectors: A sequence approach to mapping and
scanning the human genome
Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)
99380589
COMMENT
Contact: Mahairas GG, Wallace JC, Hood L
High Throughput Sequencing Center

University of Washington
401 Queen Anne Avenue North, Seattle, WA 98109, USA
Tel: (206) 616-3618
Fax: (206) 616-3887
Email: jwallace@u.washington.edu
Sequence Tagged Connector
Plate: 3080 row: G column: 8
Class: BAC ends

High quality sequence stop: 538.

FEATURES

source

Location/Qualifiers

1..538

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone_lib="CIT Approved Human Genomic Sperm Library D"

/sex="male"

/note="Organ: sperm; Vector: pBelobAC11; BAC clones in E-Coli DH10B"

BASE COUNT 125 a 131 c 117 g 163 t 2 others

ORIGIN

Query Match 83.0%; Score 82.2; DB 17; Length 538;
Best Local Similarity 91.6%; Pred. No. 5.8e-17;
Matches 87; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

OY 5 CCGTAGGCGCAAGGCAATGTTGGCATGCTGTAAGGACCACTAGATCAGACGCC 64

DB 374 CTTGAGTCTCAAGGCAATGTTGGCATGCTGTAAGGACCACTAGATCAGACGCC 315

OY 65 AGACCCCTTCTTGTGCTCAAGAAAGCGGGA 99

DB 314 GGACCCCTTCTTGTGCTCAAGAAAGCGGGA 280

RESULT 9 441 bp DNA linear GSS 09-APR-1999
B84492/c
LOCUS
DEFINITION RPEC111-2404.TP RPCI-11 Homo sapiens genomic clone RPCI-11-2404, DNA

sequence.
ACCESSION B84492
VERSION B84492.1 GI:2925624
KEYWORDS GSS.
SOURCE human.
ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 441)
Adams,M.D., Rounsley,S.D., Zhao,S., Field,C.E., Bass,S., Linher,K.,
Goiden,K., Berry,K., Granger,D., Suh,E., Wible,C., de Jong,P. and
Venter,J.C.
Use of BAC End Sequences for Sequence-Ready Map Building (1998)
Unpublished (1998)

TITLE
JOURNAL
COMMENT
Contact: Mark Adams
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: mdadams@tigr.org

Clones are derived from the human BAC library RPCI-11. For BAC
library availability, please contact Pieter de Jong
(pieter@dejong.med.buffalo.edu). Clones may be purchased from
BACPAC Resources (http://bacpac.med.buffalo.edu/ordering) or from
Research Genetics (info@resgen.com). BAC end search page:
http://www.tigr.org/tldb/humgen/bac_end_search/bac_end_search.html
Seq primer: SP6
Class: BAC ends.

FEATURES
source
Location/Qualifiers
1..441
/organism="Homo sapiens"
/db_xref="GDB:7509171"
/db_xref="taxon:9606"
/clone_lib="RPCI-11-2404"

BASE COUNT 156 a 144 c 133 g 199 t

ORIGIN
Query Match 82.2%; Score 81.4; DB 17; Length 632;
Best Local Similarity 88.9%; Pred. No. 1.1e-16;
Matches 88; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

OY 1 TCACCCCTGAGCGCAAGGCAATGTTGGCATGCTGTAAGGACCACTAGATCAGACA 60

/clone_lib="RPCI-11"
/sex="Male"
/cell_type="Lymphocytes"
/note="Vector: pBAC3.6; Site_1: EcoRI; Site_2: EcoRI;
RPEC111 Human Male BAC library"

BASE COUNT 106 a 108 c 94 g 133 t

Query Match 82.2%; Score 81.4; DB 17; Length 441;
Best Local Similarity 88.9%; Pred. No. 1e-16;
Matches 88; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

OY 1 TCACCCCTGAGCGCAAGGCAATGTTGGCATGCTGTAAGGACCACTAGATCAGACA 60

DB 307 TCACCTCTGACTGAAAGGCAATGTTGAAAGCTGGTAAGGACCACTAGATCAGACA 248

OY 61 GCCCAGACCCCTTCTTGTGCTCAAGAAAGCGGGA 99

DB 247 ACCCAGACCCCTTCTTGTGCTCAAGAAAGGGA 209

RESULT 10 632 bp DNA linear GSS 14-APR-1999
A0014046/c
LOCUS
DEFINITION RPEC111-2404.TKBR RPCI-11 Homo sapiens genomic clone RPCI-11-2404,
DNA sequence.

ACCESSION A0014046
VERSION A0014046.1 GI:3186611
KEYWORDS GSS.
SOURCE human.
ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 632)
Adams,M.D., Rounsley,S.D., Zhao,S., Field,C.E., Bass,S., Linher,K.,
Goiden,K., Berry,K., Granger,D., Suh,E., Wible,C., de Jong,P. and
Venter,J.C.
Use of BAC End Sequences for Sequence-Ready Map Building (1998)
Unpublished (1998)

TITLE
JOURNAL
COMMENT
Contact: Mark Adams
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: mdadams@tigr.org

Clones are derived from the human BAC library RPCI-11. For BAC
library availability, please contact Pieter de Jong
(pieter@dejong.med.buffalo.edu). Clones may be purchased from
BACPAC Resources (http://bacpac.med.buffalo.edu/ordering) or from
Research Genetics (info@resgen.com). BAC end search page:
http://www.tigr.org/tldb/humgen/bac_end_search/bac_end_search.html
Class: BAC ends.

FEATURES
source
Location/Qualifiers
1..632
/organism="Homo sapiens"
/db_xref="GDB:7509171"
/db_xref="taxon:9606"
/clone_lib="RPCI-11-2404"

BASE COUNT 156 a 144 c 133 g 199 t

ORIGIN
Query Match 82.2%; Score 81.4; DB 17; Length 632;
Best Local Similarity 88.9%; Pred. No. 1.1e-16;
Matches 88; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

OY 1 TCACCCCTGAGCGCAAGGCAATGTTGGCATGCTGTAAGGACCACTAGATCAGACA 60

DB 321 TCACCTCTGACCTGAAGCAATGTGATGATCTGTAGGACACATGATCCAGCA 262
OY 61 GCCACAGCCCTTCTTGTGTGTCAGAAAGCGGAAA 99
DB 261 ACCCAGACCCCTTCTTGTGTGTCAGAAAGAGGAAA 223

RESULT 11
LOCUS AO60950 624 bp DNA linear GSS 31-JUL-1998
DEFINITION CIT-HSP-2347H19.TF CIT-HSP Homo sapiens genomic clone 2347H19, DNA sequence.
ACCESSION AO60950
VERSION AO60950.1 GI:3362862
KEYWORDS GSS.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 624)
Adams M.D., Rounsley S.D., Zhao S., Field C.E., Bass S., Linher K., Golden K., Berry K., Gauger D., Sun E., Wible C., Shizuya H., Simon M. and Venter J.C.
Use of a random BAC End Sequence Database for Sequence-Ready Map Building (1998)
Unpublished (1998)
OTHER GSSs: CIT-HSP-2347H19.TR
CONTACT: Mark Adams
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: mdadams@tigr.org
Clones are available from Research Genetics (info@resgen.com). BAC end search page:
http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html
Seq primer: M13-21
Class: BAC ends

FEATURES
Location/Qualifiers
source 1..624
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="2347H19"
/clone_1lb="CIT-HSP"
/sex="Male"
/cell_type="Sperm"
/note="Vector: pBelobAC11, site_1: HindIII, site_2: HindIII"
BASE COUNT 244 a 130 c 125 g 125 t
ORIGIN

Query Match 54.1%; Score 53.6; DB 17; Length 624;
Best Local Similarity 93.3%; Pred. No. 2.2e-07;
Matches 56; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

OY 40 AAGCAGCACTAGATCCAGAGCCAGCCCTTCTTGTGTGTCAGAAAGCGGAAA 99
DB 448 AAAAATACCTAGATCCAGAGCCAGCCCTTCTTGTGTGTCAGAAAGCGGAAA 507

RESULT 12
LOCUS AO609568 515 bp DNA linear GSS 15-JUN-1999
DEFINITION HS_5081.B1.D08.T7A RPCI-11 Human Male BAC Library Homo sapiens genomic clone Plate-657 Col-15 Row-H, DNA sequence.
ACCESSION AO609568
VERSION AO609568.1 GI:5070844
KEYWORDS GSS.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

REFERENCE 1 (bases 1 to 515)
Mahairas G.G., Wallace J.C., Smith K., Swartzell S., Holzman T., Keller A., Shaker R., Furlong J., Young J., Zhao S., Adams M.D. and Hood L.
Sequence-tagged connectors: A sequence approach to mapping and scanning the human genome
Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)
99380589
CONTACT: Mahairas G.G., Wallace J.C., Hood L
High Throughput Sequencing Center
University of Washington
401 Queen Anne Avenue North, Seattle, WA 98109, USA
Tel: (206) 616-3618
Fax: (206) 616-3887
Email: jwallace@u.washington.edu
Clones are derived from the human BAC library RPCI-11. For BAC library availability, please contact Pieter de Jong (pieter@dejong.med.buffalo.edu). Clones may be purchased from BACPAC Resources (http://bacpac.med.buffalo.edu/ordering_bac.htm) or from Research Genetics (info@resgen.com). BAC end Web Server: <http://www.htsc.washington.edu>
Plate: 657 row: H column: 15
Seq primer: T7
Class: BAC ends
High quality sequence stop: 515.

FEATURES
Location/Qualifiers
source 1..515
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="Plate-657 Col-15 Row-H"
/clone_1lb="RPCI-11 Human Male BAC library"
/sex="male"
/note="Vector: pBAC3.6; Site_1: EcoRI; Site_2: EcoRI; Male blood DNA was isolated from one randomly chosen donor and partially digested with a combination of EcoRI and EcoRI Methylase. Size selected DNA was cloned into the pBAC3.6 vector at EcoRI sites"
BASE COUNT 140 a 114 c 118 g 117 t 26 others
ORIGIN

Query Match 47.7%; Score 47.2; DB 17; Length 515;
Best Local Similarity 72.6%; Pred. No. 2.8e-05;
Matches 61; Conservative 0; Mismatches 23; Indels 0; Gaps 0;

OY 13 GCAAGGCAATGTGGCATGCTGTAAAGACCACTAGATCCAGAGCCAGACCCCT 72
DB 138 GCGAAGCTATAGACATGCTGTAAAGACCACTAGATCCAGAGCCAGACCCCT 197

OY 73 TTCTTGTGTCAGAAAGCGGG 96
DB 198 TTTTGTGCGCTAAAGAGTGAG 221

RESULT 13
LOCUS T96504 386 bp mRNA linear EST 27-MAR-1995
DEFINITION y34902.r1 Stratagene lung (#937210) Homo sapiens cDNA clone IMAGE:119666 5', mRNA sequence.
ACCESSION T96504
VERSION T96504.1 GI:735128
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 386)
Hillier L., Lennon G., Becker M., Bonaldo M.F., Chapell B., Chisoe S., Dietrich N., Dubuque T., Favello A., Gish W., Hawkins M., Hultman M., Kucaba T., Lacy M., Le M., Le N., Mardis E., Moore B., Morris M., Parsons J., Prange C., Rifkin L., Rohlfing T., Schellenberg K., Soares M.B., Tan F., Thierly-Meg J., Trevaaskis E., Underwood K., Wohlmann P., Waterston R., Wilson R. and Marra M.

TITLE Generation and analysis of 280,000 human expressed sequence tags
JOURNAL Genome Res. 6 (9), 807-828 (1996)
MEDLINE 97044478
COMMENT Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: estewatson.wustl.edu

Insert Size: 1072
High quality sequence stops: 254 Source: IMAGE Consortium, LNL
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Insert Length: 1072 Std Error: 0.00
Seq primer: M13RPI
High quality sequence stop: 254.

FEATURES

source Location/Qualifiers

1..386
/organism="Homo sapiens"
/db_xref="GDB:487955"
/db_xref="taxon:9606"
/clone="IMAGE:119666"
/clone_id="Stratagene lung (#937210)"
/sex="male"
/dev_stage="72 years"
/lab_host="SOLR cells (Kanamycin resistant)"
/note="Organ: Lung; Vector: pBlueScript SK-; Site_1: EcoRI
; Site_2: XhoI; Cloned unidirectionally. Primer: Oligo
dT, normal lung. Average insert size: 1.0 kb; Uni-ZAP XR
Vector; -5' adaptor sequence: 5' GAATTCGACGACGAC 3' -3'
adaptor sequence: 5' CTCGAGTCTTTTCTTTTCTTTT 3' "

BASE COUNT 108 a 83 c 102 g 90 t 3 others

ORIGIN

Query Match 34.5%; Score 34.2; DB 14; Length 386;
Best Local Similarity 92.3%; Pred. No. 0.59;
Matches 36; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 61 GCCCAGACCCTTCTTGTGTCAGAAAGCGCGGAA 99
Db 1 GCCCAGACCCTTCTTGTGTCAGAGGCGGAA 39

RESULT 14 482 bp mRNA linear EST 20-JUN-2002
LOCUS B0560733

DEFINITION H4066B08-5 NIA Mouse 7.4k cDNA clone Set Mus musculus cDNA clone
H4066B08 5', mRNA sequence.

ACCESSION B0560733

VERSION B0560733.1 GI:21461618

KEYWORDS EST

SOURCE house mouse.

ORGANISM Mus musculus

REFERENCE Eukaryota; Metazoa; Chordata; Craniala; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 482)

AUTHORS VanBuren, V., Piao, Y., Dudekula, D.B., Qian, Y., Carter, M.G., Martin
P.R., Staeg, C.A., Bassey, U., Alpa, K., Hamatani, T., Karpi, G.J.,
Luo, A.G. and Ko, M.S.H.

TITLE Assembly, verification, and initial annotation of NIA 7.4k mouse
cDNA clone set

JOURNAL Unpublished (2002)

COMMENT Contact: Yong Qian
Laboratory of Genetics
National Institute on Aging/National Institutes of Health
333 Cassell Drive, Suite 3000, Baltimore, MD 21224-6820, USA
Email: cdna@igsun.grc.nia.nih.gov
This clone set has been freely distributed to the community. Please
visit http://igsun.grc.nia.nih.gov/cDNA/NIA_7.4k.html for details.
Seq primer: -21M13 Reverse
High quality sequence stop: 482
POLYA-No.

FEATURES

source Location/Qualifiers

1..482
/organism="Mus musculus"
/strain="C57BL/6"
/db_xref="taeST:H4066B08-5"
/db_xref="taxon:10090"
/clone="H4066B08"
/clone_id="NIA Mouse 7.4k cDNA clone Set"
/sex="mixed"
/dev_stage="mixed"
/lab_host="DH10B"
/note="Vector: pSPORT1; Site_1: SalI; Site_2: NotI; This
clone is among a rearranged set of 7,407 clones from more
than 20 cDNA libraries."

BASE COUNT 117 a 124 c 126 g 115 t

ORIGIN

Query Match 31.9%; Score 31.6; DB 14; Length 482;
Best Local Similarity 62.8%; Pred. No. 4.6;
Matches 49; Conservative 0; Mismatches 29; Indels 0; Gaps 0;

OY 12 CCAGAGCAATGTTGGCATGCTGTAAAGACCACTAGATCCAGACCCAGACCC 71

Db 165 CCAGAGCAATGTTGGCATGCTGTAAAGACCACTAGATCCAGACCCAGACCC 106

OY 72 TTTCTTGTGTCACAGAA 89

Db 105 TTTCTTGTGTCACAGAA 88

RESULT 15

BM941798 683 bp mRNA linear EST 29-APR-2002
LOCUS BM941798
DEFINITION U1-M-CG0p-bq1-c-04-0-UI.r1 NIH_BMAP_Ret4_S2 Mus musculus cDNA clone
U1-M-CG0p-bq1-c-04-0-UI 5', mRNA sequence.

ACCESSION BM941798

VERSION BM941798.1 GI:19400950

KEYWORDS EST

SOURCE house mouse.

ORGANISM Mus musculus

REFERENCE Eukaryota; Metazoa; Chordata; Craniala; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 683)

AUTHORS Bonaldo, M.F., Lennon, G. and Soares, M.B.

TITLE Normalization and subtraction: two approaches to facilitate gene
discovery

JOURNAL National Institute of Mental Health
MEDLINE 6001 Executive Blvd. Room 7N-7190, MSC 9643, Bethesda, MD
20892-9643, USA
Tel: 301 443 1706
Fax: 301 443 9890
Email: mestr@mail.nih.gov
Tissue Procurement: Dr. Xin-Yuan Fu, Yale University School of
Medicine
cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Researchers may obtain clones from Research
Genetics (www.resgen.com).

Seq primer: M13 REVERSE.

Location/Qualifiers

1..683
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="U1-M-CG0p-bq1-c-04-0-UI"
/clone_id="NIH_BMAP_Ret4_S2"
/lab_host="DH10B (Life Technologies)"
/note="Vector: pRT30-Pac (Pharmacia) with a modified
polylinker. Site_1: Not I; Site_2: Eco RI; The

GenCore version 5.1.5
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 2, 2003, 11:30:30 ; Search time 39.6 Seconds

(without alignments)
766,692 Million cell updates/sec

Title: US-09-719-554-3_COPY_2502_2600

Perfect score: 99
Sequence: 1 tcaacctgagcgcaaacg.....tgctcaagaagcggaagaa 99

Scoring table: IDENTITY_NUC
Gapop 10.0 ; Gapext 1.0

Searched: 441362 seqs, 15338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents NA: *
1: /cgn2_6/ptodata/2/ina/5A.COMB.seq: *
2: /cgn2_6/ptodata/2/ina/6A.COMB.seq: *
3: /cgn2_6/ptodata/2/ina/6A.COMB.seq: *
4: /cgn2_6/ptodata/2/ina/6B.COMB.seq: *
5: /cgn2_6/ptodata/2/ina/PCrus.COMB.seq: *
6: /cgn2_6/ptodata/2/ina/Backflist.seq: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	31.8	32.1	1100	4	US-08-456-200B-17
2	31.8	32.1	2470	1	US-07-745-206A-14
3	31.8	32.1	2470	2	US-08-311-363-14
4	31.8	32.1	5467	1	US-07-745-206A-12
5	31.8	32.1	5467	2	US-08-311-363-12
6	31.8	32.1	7175	1	US-08-455-543A-8
7	31.8	32.1	7175	2	US-08-193-078B-8
8	31.8	32.1	7175	2	US-08-223-305C-8
9	31.8	32.1	7175	2	US-08-149-097D-8
10	31.8	32.1	7175	3	US-08-948-386-8
11	31.8	32.1	7175	3	US-08-450-562-8
12	31.8	32.1	7175	4	US-08-984-709A-8
13	31.8	32.1	7175	4	US-08-450-272-8
14	31.8	32.1	7177	4	US-09-268-163-7
15	31.8	32.1	7266	3	US-08-713-118-1
16	31.8	32.1	7266	3	US-09-452-007-1
17	31.8	32.1	7362	1	US-08-455-543A-7
18	31.8	32.1	7362	2	US-08-193-078B-7
19	31.8	32.1	7362	2	US-08-223-305C-7
20	31.8	32.1	7362	3	US-08-149-097D-7
21	31.8	32.1	7362	3	US-08-948-386-7
22	31.8	32.1	7362	4	US-08-450-562-7
23	31.8	32.1	7362	4	US-08-984-709A-7
24	31.8	32.1	7362	4	US-08-450-272-7
25	31.8	32.1	7364	4	US-09-268-163-5
26	31.8	32.1	7376	4	US-09-268-163-3
27	30.3	30.3	7011	4	US-09-268-163-9

28	26.8	27.1	4403765	4	US-09-103-840A-2	Sequence 2, Appl
29	26.8	27.1	4411529	4	US-09-103-840A-1	Sequence 1, Appl
30	26.6	26.9	590	1	US-08-580-038-8	Sequence 8, Appl
31	26.4	26.7	6038	4	US-09-305-639-4	Sequence 4, Appl
32	26.4	26.7	7622	4	US-09-305-639-1	Sequence 1, Appl
33	26.3	26.3	3117	3	US-08-909-954-3	Sequence 3, Appl
34	25.8	26.1	4136	4	US-09-103-875-2	Sequence 2, Appl
35	25.6	25.9	163450	4	US-09-345-882-1	Sequence 1, Appl
36	25.4	25.7	2301	4	US-09-448-218D-1	Sequence 1, Appl
37	25.4	25.7	2301	4	US-09-448-218D-3	Sequence 3, Appl
38	25.4	25.7	2301	4	US-09-449-218D-5	Sequence 5, Appl
39	25.4	25.7	2301	4	US-09-449-218D-7	Sequence 7, Appl
40	25.4	25.7	9301	4	US-09-449-218D-18	Sequence 18, Appl
41	25.2	25.5	138	3	US-08-812-121-17	Sequence 17, Appl
42	25.2	25.5	138	4	US-09-403-672-17	Sequence 17, Appl
43	25	25.3	595	1	US-08-580-038-25	Sequence 25, Appl
44	25	25.3	2004	1	US-08-471-033-18	Sequence 18, Appl
45	25	25.3	2004	2	US-08-471-044-18	Sequence 18, Appl

ALIGNMENTS

RESULT 1
US-08-456-200B-17
Sequence 17, Application US/08456200B
Patent No. 6229000
GENERAL INFORMATION:
APPLICANT: Franz, Jurgen; Weingartner, Bernhard;
APPLICANT: Unterbeck, Axel; Rae, Peter
TITLE OF INVENTION: TISSUE-SPECIFIC HUMAN NEURONL
TITLE OF INVENTION: CALCIUM CHANNEL SUB-TYPES AND
THEIR USE
NUMBER OF SEQUENCES: 17
CORRESPONDENCE ADDRESSES:
ADDRESSEE: SPRUNG HORN KRAMER & WOODS
STREET: 660 White Plains Road
CITY: Tarrytown
STATE: New York
COUNTRY: U.S.A.
ZIP: 10591-5144
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.5 inch, 1.44 MB
MEDIUM TYPE: Storage
COMPUTER: NEC Powermate SX/20
OPERATING SYSTEM: DOS
SOFTWARE: Wordperfect 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/456,200B
FILING DATE: 31-MAY-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/094,712
FILING DATE: 19-JUL-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/858,278
FILING DATE: 26-MAR-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/064,778
FILING DATE: 19-MAY-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: DE 41 10 785
FILING DATE: 04-APR-1991
ATTORNEY/AGENT INFORMATION:
NAME: Kurt G. Briscoe
REGISTRATION NUMBER: 33, 141
REFERENCE/DOCKET NUMBER: Bayer 8398.3-KGB
TELECOMMUNICATION INFORMATION:
TELEPHONE: (914) 332-1700
TELEFAX: (914) 332-1844
INFORMATION FOR SEQ ID NO: 17:
SEQUENCE CHARACTERISTICS:

LENGTH: 1100 nucleotides
TYPE: Nucleotide
STRANDEDNESS: Single
TOPOLOGY: Linear
MOLECULE TYPE: cDNA
US-08-456-200B-17

Query Match 32.1%; Score 31.8; DB 4; Length 1100;
Best Local Similarity 61.4%; Pred. No. 0.015;
Matches 51; Conservative 0; Mismatches 32; Indels 0; Gaps 0;

QY 9 GAGCGCAAGGCAATGTTGGCATGCTGTAAGAGACCACTAGATCCAGAGCCAGAC 68
||| ||||| | ||||| ||| ||| ||| |||
DB 340 GATCTTCAAGCGGAGCAATGCTGCGCCGAGAGAGACAGATGACAGAGAGAGTC 399

OY 69 CCCTTCTTGTGTGTCAGAAAG 91
||||| ||| ||||| |||
DB 400 CCCTTGGACGTGCTGAAGAGAG 422

RESULT 2

US-07-745-206A-14
; Sequence 14, Application US/07745206A
; Patent No. 5428921
; GENERAL INFORMATION:
; APPLICANT: Harpold, Michael
; APPLICANT: Ellis, Steven
; APPLICANT: Williams, Mark
; APPLICANT: McCue, Ann
; APPLICANT: Feldman, Daniel
; TITLE OF INVENTION: Human Calcium Channel Compositions and
; TITLE OF INVENTION: Methods
; NUMBER OF SEQUENCES: 32
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fitch, Even, Tabin & Flannery
; STREET: 135 S. LaSalle
; CITY: Chicago
; STATE: Illinois
; COUNTRY: U.S.A.
; ZIP: 60603
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07745,206A
; FILING DATE: 19910815
; CLASSIFICATION: 435.
; ATTORNEY/AGENT INFORMATION:
; NAME: Feder, Scott B
; REFERENCE/DOCKET NUMBER: 51504
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-372-7842
; INFORMATION FOR SEQ. ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2470 base pairs
; TYPE: NUCLEIC ACID
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..2469
US-07-745-206A-14

Query Match 32.1%; Score 31.8; DB 1; Length 2470;
Best Local Similarity 61.4%; Pred. No. 0.021;
Matches 51; Conservative 0; Mismatches 32; Indels 0; Gaps 0;

OY 9 GAGCGCAAGGCAATGTTGGCATGCTGTAAGAGACCACTAGATCCAGAGCCAGAC 68
||| ||||| | ||||| ||| ||| ||| |||
DB 1173 GATCTTCAAGCGGAGCAATGCTGCGCCGAGAGAGACAGATGACAGAGAGAGTC 1232

OY 69 CCCTTCTTGTGTGTCAGAAAG 91
||||| ||| ||||| |||
DB 1233 CCCTTGGACGTGCTGAAGAGAG 1255

RESULT 3

US-08-311-363-14
; Sequence 14, Application US/08311363
; Patent No. 5876958
; GENERAL INFORMATION:
; APPLICANT: Harpold, Michael
; APPLICANT: Ellis, Steven
; APPLICANT: Williams, Mark
; APPLICANT: Feldman, Daniel
; APPLICANT: McCue, Ann
; APPLICANT: Brenner, Robert
; TITLE OF INVENTION: Human Calcium Channel Compositions and
; TITLE OF INVENTION: Methods
; NUMBER OF SEQUENCES: 32
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Brown, Martin, Haller & McClain
; STREET: 1660 Union Street
; CITY: San Diego
; STATE: California
; COUNTRY: USA
; ZIP: 92101-2926
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/311,363
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/745,206
; FILING DATE: 15-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Seidman, Stephanie L.
; REGISTRATION NUMBER: 33,779
; REFERENCE/DOCKET NUMBER: 6362-51506
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619)238-0999
; TELEFAX: (619)238-0062
; INFORMATION FOR SEQ. ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2470 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..2469
US-08-311-363-14

Query Match 32.1%; Score 31.8; DB 2; Length 2470;
Best Local Similarity 61.4%; Pred. No. 0.021;
Matches 51; Conservative 0; Mismatches 32; Indels 0; Gaps 0;

OY 9 GAGCGCAAGGCAATGTTGGCATGCTGTAAGAGACCACTAGATCCAGAGCCAGAC 68
||| ||||| | ||||| ||| ||| ||| |||
DB 1173 GATCTTCAAGCGGAGCAATGCTGCGCCGAGAGAGACAGATGACAGAGAGAGTC 1232
OY 69 CCCTTCTTGTGTGTCAGAAAG 91
||||| ||| ||||| |||
DB 1233 CCCTTGGACGTGCTGAAGAGAG 1255

RESULT 4

US-07-745-206A-12
; Sequence 12, Application US/07745206A

Patent No. 5429921
GENERAL INFORMATION:
APPLICANT: Harpold, Michael
APPLICANT: Ellis, Steven
APPLICANT: Williams, Mark
APPLICANT: McCue, Ann
TITLE OF INVENTION: Human Calcium Channel Compositions and
METHODS
NUMBER OF SEQUENCES: 32
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fitch, Even, Tabin & Flannery
STREET: 135 S. LaSalle
CITY: Chicago
STATE: Illinois
COUNTRY: U.S.A.
ZIP: 60603
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/745,206A
FILING DATE: 19910815
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Feder, Scott B
REFERENCE/DOCKET NUMBER: 51504
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312-372-7842
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 5467 base pairs
TYPE: NUCLEIC ACID
STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: join(144..3164, 3168..3245, 3249..3386, 3390
LOCATION: ..3392, 3396..3488, 3495..3539, 3543..3581, 3585
LOCATION: ..3587, 3591..3626, 3630..3689, 3693..3737, 3744
LOCATION: ..3746, 3750..4823, 4827..4841, 4845..5006, 5010
LOCATION: ..5096, 5100..5306, 5310..5366, 5370..5465)
US-07-745-206A-12
Query Match 32.1%; Score 31.8; DB 1; Length 5467;
Best Local Similarity 61.4%; Pred. No. 0.03;
Matches 51: Conservative 0; Mismatches 32; Indels 0; Gaps 0;
QY 9 GAGCCCAAGCAATGTTGGCATCTGCTAAAGACACCTAGATCCAGACCCAGAC 68
DB 1316 GATCTTCAAGCGGAGGAAGTATGTCGCCGAGAGGACAGAGATGACAGAGAATGC 1375
QY 69 CCCTTCTTGTGTCAGAAAG 91
DB 1376 CCCTTGGACGTCTGAAGAGAG 1398
RESULT 5
US-08-311-363-12
Sequence 12, Application US/08311363
Patent No. 587658
GENERAL INFORMATION:
APPLICANT: Harpold, Michael
APPLICANT: Ellis, Steven
APPLICANT: Williams, Mark
APPLICANT: Feldman, Daniel
APPLICANT: McCue, Ann
APPLICANT: Brenner, Robert
TITLE OF INVENTION: Human Calcium Channel Compositions and
METHODS

NUMBER OF SEQUENCES: 32
CORRESPONDENCE ADDRESS:
ADDRESSEE: Brown, Martin, Haller & McClain
STREET: 1660 Union Street
CITY: San Diego
STATE: California
COUNTRY: USA
ZIP: 92101-2926
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/311,363
FILING DATE:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/745,206
FILING DATE: 15-AUG-1991
ATTORNEY/AGENT INFORMATION:
NAME: Seidman, Stephanie L.
REGISTRATION NUMBER: 33,779
REFERENCE/DOCKET NUMBER: 6362-51506
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619)238-0999
TELEFAX: (619)238-0062
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 5467 base pairs
TYPE: nucleic acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: join(144..3164, 3168..3245, 3249..3386, 3390
LOCATION: ..3392, 3396..3488, 3495..3539, 3543..3581, 3585
LOCATION: ..3587, 3591..3626, 3630..3689, 3693..3737, 3744
LOCATION: ..3746, 3750..4823, 4827..4841, 4845..5006, 5010
LOCATION: ..5096, 5100..5306, 5310..5366, 5370..5465)
US-08-311-363-12
Query Match 32.1%; Score 31.8; DB 2; Length 5467;
Best Local Similarity 61.4%; Pred. No. 0.03;
Matches 51: Conservative 0; Mismatches 32; Indels 0; Gaps 0;
QY 9 GAGCCCAAGCAATGTTGGCATCTGCTAAAGACACCTAGATCCAGACCCAGAC 68
DB 1316 GATCTTCAAGCGGAGGAAGTATGTCGCCGAGAGGACAGAGATGACAGAGAATGC 1375
QY 69 CCCTTCTTGTGTCAGAAAG 91
DB 1376 CCCTTGGACGTCTGAAGAGAG 1398
RESULT 6
US-08-455-543A-8
Sequence 8, Application US/08455543A
Patent No. 5792846
GENERAL INFORMATION:
APPLICANT: Harpold, Michael
APPLICANT: Ellis, Steven
APPLICANT: Williams, Mark
APPLICANT: Feldman, Daniel
APPLICANT: McCue, Ann
APPLICANT: Brenner, Robert
TITLE OF INVENTION: HUMAN CALCIUM CHANNEL COMPOSITIONS AND
METHODS
NUMBER OF SEQUENCES: 57
CORRESPONDENCE ADDRESS:
ADDRESSEE: Brown, Martin, Haller & McClain
STREET: 1660 Union Street
CITY: San Diego

```

1 STATE: California
2 COUNTRY: USA
3 ZIP: 92101-2926
4
5 COMPUTER READABLE FORM:
6 MEDIUM TYPE: Diskette
7 OPERATING SYSTEM: DOS
8 SOFTWARE: Fastseq Version 1.5
9 CURRENT APPLICATION DATA:
10 APPLICATION NUMBER: US/08/455,543A
11 FILING DATE: May 31, 1995
12
13 PRIOR APPLICATION DATA:
14 APPLICATION NUMBER: 08/223,305
15 FILING DATE: April 4, 1994
16
17 PRIOR APPLICATION DATA:
18 APPLICATION NUMBER: 07/868,354
19 FILING DATE: April 10, 1992
20
21 PRIOR APPLICATION DATA:
22 APPLICATION NUMBER: US 07/745,206
23 FILING DATE: 15-AUG-1991
24
25 PRIOR APPLICATION DATA:
26 APPLICATION NUMBER: US 07/620,250
27 FILING DATE: 30-NOV-1990
28
29 PRIOR APPLICATION DATA:
30 APPLICATION NUMBER: US 07/482,384
31 FILING DATE: 20-FEB-1990
32
33 PRIOR APPLICATION DATA:
34 APPLICATION NUMBER: US 07/603,751
35 FILING DATE: 04-APR-1989
36
37 PRIOR APPLICATION DATA:
38 APPLICATION NUMBER: WO PCT/US89/01408
39 FILING DATE: 04-APR-1989
40
41 PRIOR APPLICATION DATA:
42 APPLICATION NUMBER: US 07/176,899
43 FILING DATE: 04-APR-1988
44
45 ATTORNEY/AGENT INFORMATION:
46 NAME: Seliman, Stephanie L.
47 REGISTRATION NUMBER: 33,779
48
49 REFERENCE/DOCKET NUMBER: 6362-52517
50 TELECOMMUNICATION INFORMATION:
51 TELEPHONE: (619)238-0999
52 TELEFAX: (619)238-0062
53
54 INFORMATION FOR SEQ ID NO: 8:
55
56 SEQUENCE CHARACTERISTICS:
57 LENGTH: 7175 base pairs
58 TYPE: nucleic acid
59 STRANDEDNESS: double
60 TOPOLOGY: linear
61
62 MOLECULE TYPE: DNA (genomic)
63
64 FEATURE:
65 NAME/KEY: CDS
66 LOCATION: 144..6857
67
68 FEATURE:
69 NAME/KEY: 5'UTR
70 LOCATION: 1..143
71
72 FEATURE:
73 NAME/KEY: 3'UTR
74 LOCATION: 6855..7175
75
76 US-08-455-543A-8

```

[illegible]

```

RESULT 7
US-08-193-078B-8
: Sequence 8, Application US/08193078B
: Patent No. 5846757
: GENERAL INFORMATION:
: APPLICANT: Harpold, Michael
: APPLICANT: Ellis, Steven
: APPLICANT: Williams, Mark
: APPLICANT: Feldman, Daniel
: APPLICANT: McCue, Ann
: APPLICANT: Brenner, Robert
: TITLE OF INVENTION: HUMAN CALCIUM CHANNEL COMPOSITIONS AND
: TITLE OF INVENTION: METHODS
: NUMBER OF SEQUENCES: 29
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: BROWN, MARTIN, HALLER & MCCLAIN
: STREET: 1660 UNION STREET
: CITY: SAN DIEGO
: STATE: CA
: COUNTRY: USA
: ZIP: 92101
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patent in Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/193,078B
: FILING DATE: 07-FEB-1994
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 07/868,354
: FILING DATE: 10-APR-1992
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 07/745,206
: FILING DATE: 15-AUG-1991
: ATTORNEY/AGENT INFORMATION:
: NAME: Seidman, Stephanie L.
: REGISTRATION NUMBER: 33,779
: REFERENCE/DOCKET NUMBER: 6362-53607
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 619-238-0062
: TELEFAX: 619-238-0999
: INFORMATION FOR SEQ ID NO: 8:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 7175 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: double
: TOPOLOGY: linear
: MOLECULE TYPE: DNA (genomic)
: FEATURE:
: NAME/KEY: CDS
: LOCATION: 144..6857
: FEATURE:
: NAME/KEY: 5'UTR
: LOCATION: 1..143
: FEATURE:
: NAME/KEY: 3'UTR
: LOCATION: 6855..7175
: US-08-193-078B-8

```

	Query Match	32.1%	Score 31.8;	DB 2,	Length 7175;	
	Best Local Similarity	61.4%;	Pred. No. 0.033;	Mismatches 32;	Gaps 0;	
	Matches	51;	Conservative	0;		
OY	9	GAGCGCAAAAGC	CAATGTTGGCATCTGTGAAGCACACATGAA	TCCAGACGCCAGAC	68	
Dd	1316	GATCTTCAAGCGGAGGAGA	GCATCTTGTCGGCCGAGGAGGACAGAA	TGCAGAGAAATGCAGAGAGAA	1375	
OY	69	CCCTTCTTTG	TGNGTCAGAAAG	91		
Dd	1376	CCCTTGGACGTGCTGA	AAGAGAG	1398		

RESULT 8
US-08-223-305C-8
Sequence 8, Application US/08223305C
Patent No. 5851824
GENERAL INFORMATION:
APPLICANT: Harpold, Michael
APPLICANT: Ellis, Steven
APPLICANT: Williams, Mark
APPLICANT: Feldman, Daniel
APPLICANT: McCue, Ann
APPLICANT: Brenner, Robert
TITLE OF INVENTION: HUMAN CALCIUM CHANNEL COMPOSITIONS AND
METHODS
NUMBER OF SEQUENCES: 57
CORRESPONDENCE ADDRESS:
ADDRESSEE: Brown, Martin, Haller & McClain
STREET: 1660 Union Street
CITY: San Diego
STATE: California
COUNTRY: USA
ZIP: 92101-2926
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ Version 1.5
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/223,305C
FILING DATE: April 4, 1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/868,354
FILING DATE: April 10, 1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/745,206
FILING DATE: 15-AUG-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/620,250
FILING DATE: 30-NOV-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/482,384
FILING DATE: 20-FEB-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/603,751
FILING DATE: 04-APR-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO PCT/US89/01408
FILING DATE: 04-APR-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/176,899
FILING DATE: 04-APR-1988
ATTORNEY/AGENT INFORMATION:
NAME: Seidman, Stephanie L.
REGISTRATION NUMBER: 33,779
REFERENCE/DOCKET NUMBER: 52516 (P519739)
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619)238-0999
TELEFAX: (619)238-0062
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 7175 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 144..6857
NAME/KEY: 5'UTR
LOCATION: 1..143
FEATURE:
NAME/KEY: 3'UTR

LOCATION: 6855..7175
US-08-223-305C-8
Query Match 32.1%; Score 31.8; DB 2; Length 7175;
Best Local Similarity 61.4%; Pred. No. 0.033;
Matches 51; Conservative 0; Mismatches 32; Indels 0; Gaps 0;
OY 9 GAGCCGAAGCAATGTTGGGATCTGTAAAGACCACTAGATCCAGCAGCCAGAC 68
||| ||||| | | ||||| ||||| ||| ||| ||| |||
DB 1316 GATCTTCAAGGCGGAGGAGTATCTGCCGAGAGACAGCAATGCAGAGAGATC 1375
OY 69 CCCCTTCTTGTGTCAGAAAG 91
||||| ||| ||||| ||
DB 1376 CCCTTGACGCTGTGAAGAGAG 1398
RESULT 9
US-08-149-097D-8
Sequence 8, Application US/08149097D
Patent No. 5874236
GENERAL INFORMATION:
APPLICANT: Harpold, Michael
APPLICANT: Ellis, Steven
APPLICANT: Williams, Mark
APPLICANT: Feldman, Daniel
APPLICANT: McCue, Ann
APPLICANT: Brenner, Robert
TITLE OF INVENTION: HUMAN CALCIUM CHANNEL COMPOSITIONS AND
METHODS
NUMBER OF SEQUENCES: 40
CORRESPONDENCE ADDRESS:
ADDRESSEE: Brown, Martin, Haller & McClain
STREET: 1660 Union Street
CITY: San Diego
STATE: California
COUNTRY: USA
ZIP: 92101-2926
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/149,097D
FILING DATE: 05-NOV-1993
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/105,536
FILING DATE: 11-AUG-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO PCT/US92/06903
FILING DATE: 14-AUG-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/914,231
FILING DATE: 13-JUL-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/868,354
FILING DATE: 10-APR-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/745,206
FILING DATE: 15-AUG-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/620,250
FILING DATE: 30-NOV-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/482,384
FILING DATE: 20-FEB-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/603,751
FILING DATE: 04-APR-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO PCT/US89/01408
FILING DATE: 04-APR-1989

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/176,899
FILING DATE: 04-APR-1988
ATTORNEY/AGENT INFORMATION:
NAME: Seidman, Stephanie L.
REGISTRATION NUMBER: 33,779
REFERENCE/DOCKET NUMBER: 6362-55038
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 238-0999
TELEFAX: (619) 238-0062
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 7175 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 144..6857
FEATURE:
NAME/KEY: 5'UTR
LOCATION: 1..143
FEATURE:
NAME/KEY: 3'UTR
LOCATION: 6855..7175
US-08-149-097D-8

Query Match 32.1%; Score 31.8; DB 2; Length 7175;
Best Local Similarity 61.4%; Pred. No. 0.033;
Matches 51; Conservative 0; Mismatches 32; Indels 0; Gaps 0;
DB 1316 GATCTTCAGAGCGGAGAGTCTGCTGCGCGGAGAGACGAGATGCAAGAGAGAGTCTC 1375
QY 69 CCCTTCTTTGTGTCAGAGAAG 91
DB 1376 CCCTTGGACGTGCTGTAAGAGAG 1398

RESULT 10
US-08-949-386-8
Sequence 8, Application US/08949386
Patent No. 6090623
GENERAL INFORMATION:
APPLICANT: Harpold, Michael
APPLICANT: Ellis, Steven
APPLICANT: Williams, Mark
APPLICANT: McCue, Ann
APPLICANT: Gillespie, Allison
TITLE OF INVENTION: HUMAN CALCIUM CHANNEL COMPOSITIONS AND
METHODS
NUMBER OF SEQUENCES: 38
CORRESPONDENCE ADDRESS:
ADDRESSEE: Brown, Martin, Haller & McClain
STREET: 1660 Union Street
CITY: San Diego
STATE: California
COUNTRY: US
ZIP: 92101
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/290,012
FILING DATE: 11-AUG-1994

APPLICATION NUMBER: 08/149,097
FILING DATE: 5-NOV-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/105,536
FILING DATE: 11-AUG-1993
ATTORNEY/AGENT INFORMATION:
NAME: Seidman, Stephanie L.
REGISTRATION NUMBER: 33,779
REFERENCE/DOCKET NUMBER: 519808
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 238-0999
TELEFAX: (619) 238-0062
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 7175 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 144..6857
FEATURE:
NAME/KEY: 5'UTR
LOCATION: 1..143
FEATURE:
NAME/KEY: 3'UTR
LOCATION: 6855..7175
US-08-949-386-8

Query Match 32.1%; Score 31.8; DB 3; Length 7175;
Best Local Similarity 61.4%; Pred. No. 0.033;
Matches 51; Conservative 0; Mismatches 32; Indels 0; Gaps 0;
DB 1316 GATCTTCAGAGCGGAGAGTCTGCTGCGCGGAGAGACGAGATGCAAGAGAGAGTCTC 1375
QY 69 CCCTTCTTTGTGTCAGAGAAG 91
DB 1376 CCCTTGGACGTGCTGTAAGAGAG 1398

RESULT 11
US-08-450-562-8
Sequence 8, Application US/08450562
Patent No. 6096514
GENERAL INFORMATION:
APPLICANT: Harpold, Michael
APPLICANT: Ellis, Steven
APPLICANT: Williams, Mark
APPLICANT: McCue, Ann
APPLICANT: Gillespie, Allison
APPLICANT: Feldman, Daniel
APPLICANT: Brenner, Robert
TITLE OF INVENTION: HUMAN CALCIUM CHANNEL COMPOSITIONS AND
METHODS
NUMBER OF SEQUENCES: 38
CORRESPONDENCE ADDRESS:
ADDRESSEE: Brown, Martin, Haller & McClain
STREET: 1660 Union Street
CITY: San Diego
STATE: California
COUNTRY: US
ZIP: 92101
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/450,562
FILING DATE:

CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/404,950
FILING DATE: 13-MAR-1995
APPLICATION NUMBER: 08/336,257
FILING DATE: 7-NOV-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/314,083
FILING DATE: 28-SEPT-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/311,363
FILING DATE: 23-SEPT-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/290,012
FILING DATE: 11-AUG-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/223,305
FILING DATE: 4-APR-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/193,078
FILING DATE: 07-FEB-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/149,097
FILING DATE: 5-NOV-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/105,536
FILING DATE: 11-AUG-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/914,231
FILING DATE: 13-JULY-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/868,354
FILING DATE: 10-APR-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US92/06903
FILING DATE: 14-AUG-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/745,206
FILING DATE: 15-AUG-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/620,250
FILING DATE: 30-NOV-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/603,751
FILING DATE: 08-NOV-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/482,384
FILING DATE: 02-FEB-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US89/01408
FILING DATE: 04-APR-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/176,899
FILING DATE: 04-APR-1988
ATTORNEY/AGENT INFORMATION:
NAME: Seidman, Stephanie L.
REGISTRATION NUMBER: 33,779
REFERENCE/DOCKET NUMBER: 6362-519812
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 238-0999
TELEFAX: (619) 238-0062
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 7175 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 144..6857
NAME/KEY: 5'UTR

LOCATION: 1..143
FEATURE:
NAME/KEY: 3'UTR
LOCATION: 6855..7175
US-08-450-562-8

Query Match 32.1%; Score 31.8; DB 3; Length 7175;
Best Local Similarity 61.4%; Pred. No. 0.033;
Matches 51; Conservative 0; Mismatches 32; Indels 0; Gaps 0;

QY 9 GAGCGCAAGCAATGTGGCATGCTGTAAGACACATAGATCCAGACCCAGAC 68
DB 1316 GATCTTCAGCGCGAGAGATGCTGCGCCGAGAGACAGAGATGCAAGAGATC 1375
QY 69 CCCTTCTTTGTGTCAGAGAAG 91
DB 1376 CCCTTGGACGTGCTGAGAGAG 1398

RESULT 12
US-08-984-709A-8
Sequence 8, Application US/08984709A
Patent No. 6320032
GENERAL INFORMATION:
APPLICANT: Williams, Mark E.
APPLICANT: Staederman, Kenneth A.
APPLICANT: Harpold, Michael M.
TITLE OF INVENTION: HUMAN CALCIUM CHANNEL COMPOSITIONS AND
NUMBER OF SEQUENCES: 52
CORRESPONDENCE ADDRESS:
ADDRESSEE: Heller Ehrman White & McCauliffe
STREET: 4250 Executive Square, Suite 700
CITY: La Jolla
STATE: California
COUNTRY: US
ZIP: 92037
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq Version 1.5
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/984,709A
FILING DATE: 02-DEC-1997
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Seidman, Stephanie L.
REGISTRATION NUMBER: 33,779
REFERENCE/DOCKET NUMBER: 24735-9815 (formerly 6362-9815)
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 450-8400
TELEFAX: (619) 587-5360
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 7175 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 144..6857
FEATURE:
NAME/KEY: 5'UTR
LOCATION: 1..143
FEATURE:
NAME/KEY: 3'UTR
LOCATION: 6855..7175
US-08-984-709A-8
Query Match 32.1%; Score 31.8; DB 4; Length 7175;
Best Local Similarity 61.4%; Pred. No. 0.033;

```
Matches 51; Conservative 0; Mismatches 32; Indels 0; Gaps 0;
OY 9 GAGCGCAAGCATGTGGCATCTGTAAAGACCACTAGATCCAGAGCCCGAGAC 68
    ||| ||| | ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1316 GATCTTAAGCGGAGAGTCAATGCTGCGCGAGAGACAGATGCAGAGAGATC 1375
OY 69 CCCTTCTTGTGCTCAAGAAAG 91
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1376 CCCTTGGACGTGCTGAAGAGAG 1398

RESULT 13
US-08-450-272-8
; Sequence 8, Application US/08450272
; Patent No. 6387696
; GENERAL INFORMATION:
; APPLICANT: Harpold, Michael
; APPLICANT: Ellis, Steven
; APPLICANT: Williams, Mark
; APPLICANT: McCue, Ann
; APPLICANT: Gillespie, Allison
; APPLICANT: Feldman, Daniel
; APPLICANT: Brenner, Robert
; TITLE OF INVENTION: HUMAN CALCIUM CHANNEL COMPOSITIONS AND
; TITLE OF INVENTION: METHODS
; NUMBER OF SEQUENCES: 38
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Brown, Martin, Haller & McClain
; STREET: 1660 Union Street
; CITY: San Diego
; STATE: California
; COUNTRY: US
; ZIP: 92101
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/450,272
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/404,950
; FILING DATE: 13-MAR-1995
; APPLICATION NUMBER: 08/336,257
; FILING DATE: 7-NOV-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/314,083
; FILING DATE: 28-SEPT-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/311,363
; FILING DATE: 23-SEPT-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/290,012
; FILING DATE: 11-AUG-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/223,305
; FILING DATE: 4-APR-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/193,078
; FILING DATE: 07-FEB-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/149,097
; FILING DATE: 5-NOV-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/105,536
; FILING DATE: 11-AUG-1993
; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 07/914,231
; FILING DATE: 13-JULY-1992
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/868,354
; FILING DATE: 10-APR-1992
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US92/06903
; FILING DATE: 14-AUG-1992
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/745,206
; FILING DATE: 15-AUG-1991
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Seidman, Stephanie L.
; REGISTRATION NUMBER: 33,779
; REFERENCE/DOCKET NUMBER: 6362-519812
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 238-0999
; TELEFAX: (619) 238-0062
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7175 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 144..6857
; FEATURE:
; NAME/KEY: 5'UTR
; LOCATION: 1..143
; FEATURE:
; NAME/KEY: 3'UTR
; LOCATION: 6855..7175
; US-08-450-272-8

Query Match 32.1%; Score 31.8; DB 4; Length 7175;
Best Local Similarity 61.4%; Pred. No. 0.033;
Matches 51; Conservative 0; Mismatches 32; Indels 0; Gaps 0;
OY 9 GAGCGCAAGCATGTGGCATCTGTAAAGACCACTAGATCCAGAGCCCGAGAC 68
    ||| ||| | ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1316 GATCTTAAGCGGAGAGTCAATGCTGCGCGAGAGACAGATGCAGAGAGATC 1375
OY 69 CCCTTCTTGTGCTCAAGAAAG 91
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1376 CCCTTGGACGTGCTGAAGAGAG 1398

RESULT 14
US-09-268-163-7
; Sequence 7, Application US/09268163B
; Patent No. 6353091
; GENERAL INFORMATION:
; APPLICANT: Lipscombe, Diane
; APPLICANT: Schorge, Stephanie
; TITLE OF INVENTION: HUMAN N-TYPE CALCIUM CHANNEL ISOFORM AND USES THEREOF
; FILE REFERENCE: B1055/7000
; CURRENT APPLICATION NUMBER: US/09/268,163B
; CURRENT FILING DATE: 1999-03-12
; EARLIER APPLICATION NUMBER: US 60/077,901
; EARLIER FILING DATE: 1998-03-13
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 7
; LENGTH: 7177
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
```

NAME/KEY: CDS
LOCATION: 146..6856
US-09-268-163-7

Query Match 32.1%; Score 31.8; DB 4; Length 7177;
Best Local Similarity 61.4%; Pred. No. 0.033;
Matches 51; Conservative 0; Mismatches 32; Indels 0; Gaps 0;

QY 9 GAGGCAAGGCAATGTTGGGATGCTGTAAAGACCACTAGATCCAGCAGCCAGAC 68
DB 1318 GATCTTCAGGCGGAGGAGATGCTGTGCGGAGAGAGACAGATGCGAGAGAGAGTC 1377

QY 69 CCCTTTCTTTGGTGCAGAAAG 91
DB 1378 CCCTTTGGAGCTGCTGAAGAGAG 1400

RESULT 15

US-08-713-118-1
Sequence 1, Application US/08713118
Patent No. 6040436

GENERAL INFORMATION:

APPLICANT: Franco, Rodrigo

APPLICANT: Sun Chen, Ai Ru

TITLE OF INVENTION: NUCLEIC ACID ENCODING HUMAN NEURONAL

NUMBER OF SEQUENCES: 6

CORRESPONDENCE ADDRESSES:

ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.

STREET: Two Militia Drive

CITY: Lexington

STATE: MA

COUNTRY: USA

ZIP: 02173-4799

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/713,118

FILING DATE: 16-SEP-1996

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Mata, Elizabeth W.

REGISTRATION NUMBER: 38,236

REFERENCE/DOCKET NUMBER: ACC96-01

TELECOMMUNICATION INFORMATION:

TELEPHONE: 617-861-6240

TELEFAX: 617-861-9540

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 7266 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

MOLECULE TYPE: CDNA

FEATURE:

NAME/KEY: CDS

LOCATION: 92..7102

US-08-713-118-1

DB 1324 CCCTTTGGAGCTGCTGAAGAGAG 1346
Search completed: May 2, 2003, 12:48:32
Job time : 55.6 secs

Query Match 32.1%; Score 31.8; DB 3; Length 7266;
Best Local Similarity 61.4%; Pred. No. 0.033;
Matches 51; Conservative 0; Mismatches 32; Indels 0; Gaps 0;

QY 9 GAGGCAAGGCAATGTTGGGATGCTGTAAAGACCACTAGATCCAGCAGCCAGAC 68
DB 1264 GATCTTCAGGCGGAGGAGATGCTGTGCGGAGAGAGACAGATGCGAGAGAGAGTC 1323

QY 69 CCCTTTCTTTGGTGCAGAAAG 91
||||| ||| | |||| |

GenCore version 5.1.5
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 2, 2003, 12:25:30 ; Search time 65,862 Seconds
(without alignments)
1775.493 Million cell updates/sec

Title: US-09-719-554-3_COPY_2502_2600

Perfect score: 99
Sequence: 1 tcaccctgagcgcaagc.....tgcacgaagagcgagaa 99

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 746064 seqs, 590810554 residues

Total number of hits satisfying chosen parameters: 1492128

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_MA:*

- 1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
- 2: /cgn2_6/ptodata/1/pubpna/PCR_NEW_PUB.seq:*
- 3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
- 4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
- 5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
- 6: /cgn2_6/ptodata/1/pubpna/PCRUS_PUBCOMB.seq:*
- 7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
- 8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
- 9: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
- 10: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
- 11: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
- 12: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*
- 13: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
- 14: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	31.8	32.1	7177	12	US-10-033-026-7
2	31.8	32.1	7364	10	US-09-954-456-1179
3	31.8	32.1	7364	12	US-10-033-026-5
4	31.8	32.1	7376	12	US-10-033-026-3
5	30.3	30.3	7011	12	US-10-033-026-9
6	27.8	28.1	8794	9	US-09-764-891-8933
7	27	27.3	206	9	US-09-796-692-4050
8	27	27.3	206	9	US-10-040-862-4050
9	26.4	26.7	271	10	US-09-867-701-7757
10	26.4	26.7	6038	10	US-09-802-807-4
11	26.4	26.7	7622	10	US-09-802-807-1
12	26.2	26.5	424	9	US-09-934-531-665
13	26.2	26.5	14759	9	US-09-764-891-8344
14	26.2	26.5	30420	9	US-10-091-572-740
15	26.2	26.5	30420	9	US-09-764-891-9164
16	26.2	26.5	30420	10	US-09-764-877-2432
17	26	26.3	490	9	US-09-918-995-13251
18	26	26.3	1158	9	US-09-948-1993-2
19	26	26.3	3091	9	US-10-074-095-1001

20	26	26.3	3091	9	US-10-074-095-1002	Sequence 1002, Ap
21	26	26.3	3091	10	US-09-764-860-1001	Sequence 1001, Ap
22	26	26.3	3091	10	US-09-764-860-1002	Sequence 1002, Ap
23	25.8	26.1	167343	10	US-09-962-436-281	Sequence 281, Ap
24	25.8	26.1	167343	10	US-09-964-824A-273	Sequence 273, Ap
25	25.6	25.9	404	10	US-09-864-761-4275	Sequence 4275, Ap
26	25.6	25.9	479	9	US-09-918-995-9703	Sequence 9703, Ap
27	25.6	25.9	90541	10	US-09-759-359A-3	Sequence 3, Appl
28	25.4	25.7	245	10	US-09-878-574-13986	Sequence 13986, A
29	25.4	25.7	392	10	US-09-864-761-14440	Sequence 14440, A
30	25.4	25.7	759	10	US-09-867-274-1	Sequence 1, Appl
31	25.4	25.7	2329	9	US-10-245-103-63	Sequence 63, Appl
32	25.4	25.7	2329	9	US-10-245-107-63	Sequence 63, Appl
33	25.4	25.7	2329	9	US-10-245-143-63	Sequence 63, Appl
34	25.4	25.7	2329	9	US-10-245-171-63	Sequence 63, Appl
35	25.4	25.7	2329	9	US-10-245-851-63	Sequence 63, Appl
36	25.4	25.7	2329	9	US-10-245-883-63	Sequence 63, Appl
37	25.4	25.7	2329	9	US-10-237-535-63	Sequence 63, Appl
38	25.4	25.7	2329	9	US-10-238-183-63	Sequence 63, Appl
39	25.4	25.7	2329	9	US-10-238-183-63	Sequence 63, Appl
40	25.4	25.7	2329	9	US-10-238-370-63	Sequence 63, Appl
41	25.4	25.7	2329	9	US-10-245-055-63	Sequence 63, Appl
42	25.4	25.7	2329	9	US-10-245-147-63	Sequence 63, Appl
43	25.4	25.7	2329	9	US-10-245-730-63	Sequence 63, Appl
44	25.4	25.7	2329	9	US-10-245-739-63	Sequence 63, Appl
45	25.4	25.7	2329	9	US-10-246-210-63	Sequence 63, Appl

ALIGNMENTS

RESULT 1
US-10-033-026-7
; Sequence 7, Application US/10033026
; Patent No. US20020147309A1
; GENERAL INFORMATION:
; APPLICANT: Schorger, Stephanie
; APPLICANT: Lipscombe, Diane
; TITLE OF INVENTION: HUMAN N-TYPE CALCIUM CHANNEL ISORFORM AND USES THEREOF
; FILE REFERENCE: B1055/7000
; CURRENT APPLICATION NUMBER: US/10/033, 026
; CURRENT FILING DATE: 2001-12-28
; PRIOR APPLICATION NUMBER: US 09/268, 163
; PRIOR FILING DATE: 1999-03-12
; PRIOR APPLICATION NUMBER: US 60/077, 901
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 7
; LENGTH: 7177
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 146..6856
US-10-033-026-7

Query Match 32.1%; Score 31.8; DB 12; Length 7177;
Best Local Similarity 61.4%; Pred. No. 0.051;
Matches 51; Conservative 0; Mismatches 32; Indels 0; Gaps 0;

Db 1318 GACGCGAAGCAATGTTGGCATGCGTGAAGACACTAGATTCACAGCCAGAC 68
QY 69 CCCTTCTTGTGTGTCAGAAAG 91
Db 1378 CCCTTGAAGCTGCTGAGAGAG 1400

RESULT 2
US-09-954-456-1179
; Sequence 1179, Application US/09954456

```
Patent No. US20020115057A1
GENERAL INFORMATION:
APPLICANT: Young, Paul
TITLE OF INVENTION: Process for Identifying Anti-Cancer Therapeutic Agents Using Cand
TITLE OF INVENTION: Sets
FILE REFERENCE: 689290-76
CURRENT APPLICATION NUMBER: US/09/954,456
CURRENT FILING DATE: 2001-09-18
PRIOR APPLICATION NUMBER: US/60/233,617
PRIOR FILING DATE: 2000-09-18
PRIOR APPLICATION NUMBER: US/60/234,052
PRIOR FILING DATE: 2000-09-20
PRIOR APPLICATION NUMBER: US/60/234,923
PRIOR FILING DATE: 2000-09-25
PRIOR APPLICATION NUMBER: US/60/235,134
PRIOR FILING DATE: 2000-09-25
PRIOR APPLICATION NUMBER: US/60/235,637
PRIOR FILING DATE: 2000-09-26
PRIOR APPLICATION NUMBER: US/60/235,638
PRIOR FILING DATE: 2000-09-26
PRIOR APPLICATION NUMBER: US/60/235,711
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: US/60/235,720
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: US/60/235,840
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: US/60/235,863
PRIOR FILING DATE: 2000-09-27
NUMBER OF SEQ ID NOS: 2276
SOFTWARE: PatentIn version 3.0
SEQ ID NO 1179
LENGTH: 7364
TYPE: DNA
ORGANISM: Homo sapiens
US-09-954-456-1179

Query Match          32.1%; Score 31.8; DB 10; Length 7364;
Best Local Similarity 61.4%; Pred. No. 0.051;
Matches 51; Conservative 0; Mismatches 32; Indels 0; Gaps 0;

QY 9 GAGGCAAGGCAATGTTGGCATGCTGTAAAGACCACTAGATCCAGACCCCGAC 68
DB 1318 GATCTTCAAGCGGAGGAGATGCTGTGCGCGAGAGACAGAGATCAGAGAGAGATC 1377
QY 69 CCCTTCTTGTGCTCAAGAAG 91
DB 1378 CCCTTTGAGCTGCTGTGAAGAG 1400

RESULT 3
US-10-033-026-5
Sequence 5, Application US/10033026
Patent No. US20020147309A1
GENERAL INFORMATION:
APPLICANT: Lipscombe, Diane
APPLICANT: Schorge, Stephanie
TITLE OF INVENTION: HUMAN N-TYPE CALCIUM CHANNEL ISOFORM AND USES THEREOF
FILE REFERENCE: B1055/7000
CURRENT APPLICATION NUMBER: US/10/033,026
CURRENT FILING DATE: 2001-12-28
PRIOR APPLICATION NUMBER: US 09/268,163
PRIOR FILING DATE: 1999-03-12
PRIOR APPLICATION NUMBER: US 60/077,901
PRIOR FILING DATE: 1998-03-13
NUMBER OF SEQ ID NOS: 28
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 5
LENGTH: 7364
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: 146..7162
```

```
US-10-033-026-5
Query Match          32.1%; Score 31.8; DB 12; Length 7364;
Best Local Similarity 61.4%; Pred. No. 0.051;
Matches 51; Conservative 0; Mismatches 32; Indels 0; Gaps 0;

QY 9 GAGGCAAGGCAATGTTGGCATGCTGTAAAGACCACTAGATCCAGACCCCGAC 68
DB 1318 GATCTTCAAGCGGAGGAGATGCTGTGCGCGAGAGACAGAGATCAGAGAGAGATC 1377
QY 69 CCCTTCTTGTGCTCAAGAAG 91
DB 1378 CCCTTTGAGCTGCTGTGAAGAG 1400

RESULT 4
US-10-033-026-3
Sequence 3, Application US/10033026
Patent No. US20020147309A1
GENERAL INFORMATION:
APPLICANT: Lipscombe, Diane
APPLICANT: Schorge, Stephanie
TITLE OF INVENTION: HUMAN N-TYPE CALCIUM CHANNEL ISOFORM AND USES THEREOF
FILE REFERENCE: B1055/7000
CURRENT APPLICATION NUMBER: US/10/033,026
CURRENT FILING DATE: 2001-12-28
PRIOR APPLICATION NUMBER: US 09/268,163
PRIOR FILING DATE: 1999-03-12
PRIOR APPLICATION NUMBER: US 60/077,901
PRIOR FILING DATE: 1998-03-13
NUMBER OF SEQ ID NOS: 28
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 3
LENGTH: 7376
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: 146..7174
US-10-033-026-3

Query Match          32.1%; Score 31.8; DB 12; Length 7376;
Best Local Similarity 61.4%; Pred. No. 0.051;
Matches 51; Conservative 0; Mismatches 32; Indels 0; Gaps 0;

QY 9 GAGGCAAGGCAATGTTGGCATGCTGTAAAGACCACTAGATCCAGACCCCGAC 68
DB 1318 GATCTTCAAGCGGAGGAGATGCTGTGCGCGAGAGACAGAGATCAGAGAGAGATC 1377
QY 69 CCCTTCTTGTGCTCAAGAAG 91
DB 1378 CCCTTTGAGCTGCTGTGAAGAG 1400

RESULT 5
US-10-033-026-9
Sequence 9, Application US/10033026
Patent No. US20020147309A1
GENERAL INFORMATION:
APPLICANT: Lipscombe, Diane
APPLICANT: Schorge, Stephanie
TITLE OF INVENTION: HUMAN N-TYPE CALCIUM CHANNEL ISOFORM AND USES THEREOF
FILE REFERENCE: B1055/7000
CURRENT APPLICATION NUMBER: US/10/033,026
CURRENT FILING DATE: 2001-12-28
PRIOR APPLICATION NUMBER: US 09/268,163
PRIOR FILING DATE: 1999-03-12
PRIOR APPLICATION NUMBER: US 60/077,901
PRIOR FILING DATE: 1998-03-13
NUMBER OF SEQ ID NOS: 28
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 9
LENGTH: 7011
```

```

; TYPE: DNA
; ORGANISM: Rattus norvegicus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..7008
; US-10-033-026-9
```

```

Query Match
Best Local Similarity 30.3%; Score 30; DB 12; Length 7011;
Matches 51; Conservative 0; Mismatches 35; Indels 0; Gaps 0;
```

```

Qy 9 GAGCGCAAGCAATGTTGGCATGCTGTAAGACACACACTAGATCCAGACCCAGAC 68
Db 1173 GATCTTCAAGGGGAGGAGTATGTTGGCAGAGGACAGACAGACAGAGAGATC 1232
Qy 69 CCCTTCTTTGTGTGTCAGAAAGCG 94
Db 1233 CCCTTGTGATGCAGTGTGAAGAGAG 1258
```

RESULT 6

```

US-09-764-891-8933/c
; Sequence 8933, Application US/09764891
; Publication No. US20030077808A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC006
; CURRENT APPLICATION NUMBER: US/09/764,891
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 10231
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 8933
; LENGTH: 8794
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-09-764-891-8933
```

```

Query Match
Best Local Similarity 28.1%; Score 27.8; DB 9; Length 8794;
Matches 47; Conservative 0; Mismatches 32; Indels 0; Gaps 0;
```

```

Qy 19 GCAATGTTGGCATGCTGTAAGACACACTAGATCCAGACCCAGACCCCTTCTTT 78
Db 4260 GCTATGTTGGCCAGCTGCTTGAAGCTCTGCTCAAGCCTCCAAAGTGTGAGATT 4201
```

```

Qy 79 GTGTCAGAAAGCGCGGA 97
Db 4200 ATGGGCATGAGCCGCTGTA 4182
```

RESULT 7

```

US-09-796-692-4050/c
; Sequence 4050, Application US/09796692
; Publication No. US20020198362A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Algate, Paul A.
; APPLICANT: Mannion, Jane
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DETECTION, DIAGNOSIS AND THERAPY
; FILE REFERENCE: 2077, 001200
; CURRENT APPLICATION NUMBER: US/09/796,692
; CURRENT FILING DATE: 2001-03-01
; PRIOR APPLICATION NUMBER: 60/186,126
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: 60/190,479
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: 60/200,545
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: 60/200,303
; PRIOR FILING DATE: 2000-04-28
```

```

; PRIOR APPLICATION NUMBER: 60/200,779
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 60/200,999
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: 60/202,084
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: 60/206,201
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: 60/218,950
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: 60/222,903
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: 60/223,416
; PRIOR FILING DATE: 2000-08-04
; PRIOR APPLICATION NUMBER: 60/223,378
; PRIOR FILING DATE: 2000-08-07
; NUMBER OF SEQ ID NOS: 9597
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 4050
; LENGTH: 206
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (77)
; OTHER INFORMATION: n=A,T,C or G
; US-09-796-692-4050
```

```

Query Match
Best Local Similarity 27.3%; Score 27; DB 9; Length 206;
Matches 51; Conservative 0; Mismatches 41; Indels 0; Gaps 0;
```

```

Qy 1 TCACCCCTAGCGCAAGGCATGTTGGCATGCTGTAAGACACACTAGATCCAGCA 60
Db 160 TCAGCAGTGGCGCTGAAGACAGGCTTGAGGATGATGAGCAGCATGGGCCAGAA 101
```

```

Qy 61 GCCCAGACCCCTTCTTGTGTCAGAAAGG 92
Db 100 GCCCAGCTTCTCACCAGTAGGINTGAGAGG 69
```

RESULT 8

```

US-10-040-862-4050/c
; Sequence 4050, Application US/10040862
; Publication No. US20030078396A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Algate, Paul A.
; APPLICANT: Mannion, Jane
; APPLICANT: Retter, Marc
; APPLICANT: Corixa Corporation
; TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and The
; FILE REFERENCE: 014058-013520US
; CURRENT APPLICATION NUMBER: US/10/040,862
; CURRENT FILING DATE: 2001-11-06
; PRIOR APPLICATION NUMBER: US 60/186,126
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: US 60/190,479
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: US 60/200,545
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: US 60/200,303
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,779
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,999
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: US 60/202,084
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: US 60/206,201
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: US 60/218,950
```

;; PRIOR FILING DATE: 2000-07-14
;; PRIOR APPLICATION NUMBER: US 60/222,903
;; PRIOR FILING DATE: 2000-08-03
;; PRIOR APPLICATION NUMBER: US 60/223,416
;; PRIOR FILING DATE: 2000-08-04
;; PRIOR APPLICATION NUMBER: US 60/223,378
;; PRIOR FILING DATE: 2000-08-07
;; PRIOR APPLICATION NUMBER: US 09/796,692
;; PRIOR FILING DATE: 2001-03-01
;; NUMBER OF SEQ ID NOS: 10467
;; SOFTWARE: FastSeq for Windows Version 3.0
;; SEQ ID NO 4050
;; LENGTH: 206
;; TYPE: DNA
;; ORGANISM: Homo sapiens
;; FEATURE:
;; NAME/KEY: unsure
;; LOCATION: (77)
;; OTHER INFORMATION: n=A,T,C or G
US-10-040-862-4050

Query Match 27.3%; Score 27; DB 9; Length 206;
Best Local Similarity 55.4%; Pred. No. 1.2;
Matches 51; Conservative 0; Mismatches 41; Indels 0; Gaps 0;

OY 1 TCACCCCTGAGCGCAAGGCATGTTGGCATGCTGTAAGACCACTAGATCCAGCA 60
DB 160 TCAGAGTGGGGCTTAAGACGAGTTGGAGTGAGATGACACAGAGCCAGAA 101
OY 61 GCCCAGACCCCTTCTTTGCTGCTCAAGAAAG 92
DB 100 GCCCAGCTTCTCACCAGTGTGAGTNGAAGAG 69

RESULT 9
US-09-867-701-7757
;; Sequence 7757, Application US/09867701
;; Patent No. US20020132237A1
;; GENERAL INFORMATION:
;; APPLICANT: Aglate, Paul A.
;; APPLICANT: Jones, Robert
;; APPLICANT: Harlocker, Susan L.
;; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
;; FILE REFERENCE: 210121.497
;; CURRENT APPLICATION NUMBER: US/09/867,701
;; CURRENT FILING DATE: 2001-05-29
;; NUMBER OF SEQ ID NOS: 10912
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO 7757
;; LENGTH: 271
;; TYPE: DNA
;; ORGANISM: Homo sapiens
US-09-867-701-7757

Query Match 26.7%; Score 26.4; DB 10; Length 271;
Best Local Similarity 65.0%; Pred. No. 2.2;
Matches 39; Conservative 0; Mismatches 21; Indels 0; Gaps 0;

OY 19 GCATGTTGGCATGCTGCTAAGACCACTAGATCCAGACGCCAGCCCTTCTTT 78
DB 28 GCTATGTTGCTCAGCGCTGCTCTGACTAGACTCAAGCAATCCCTGCTGCT 87

RESULT 10
US-09-802-807-4/c
;; Sequence 4, Application US/09802807
;; Patent No. US20010034044A1
;; GENERAL INFORMATION:
;; APPLICANT: Treco, Douglas A.
;; APPLICANT: Hearle, Michael W.
;; APPLICANT: Seiden, Richard F.
;; TITLE OF INVENTION: GENOMIC SEQUENCES FOR PROTEIN PRODUCTION AND DELIVERY

;; FILE REFERENCE: 07236/016001
;; CURRENT APPLICATION NUMBER: US/09/802,807
;; CURRENT FILING DATE: 2001-03-08
;; PRIOR APPLICATION NUMBER: 60/084,663
;; PRIOR FILING DATE: 1998-05-07
;; NUMBER OF SEQ ID NOS: 7
;; SOFTWARE: FastSeq for Windows Version 3.0
;; SEQ ID NO 4
;; LENGTH: 6038
;; TYPE: DNA
;; ORGANISM: Homo sapiens
US-09-802-807-4

Query Match 26.7%; Score 26.4; DB 10; Length 6038;
Best Local Similarity 65.0%; Pred. No. 5.4;
Matches 39; Conservative 0; Mismatches 21; Indels 0; Gaps 0;

OY 19 GCATGTTGGCATGCTGCTAAGACCACTAGATCCAGACGCCAGCCCTTCTTT 78
DB 174 GCATGTTGGCGCAGCTGCTCTCAAAATCCTAGACTCAAGCAATCCACCTTGGCTT 115

RESULT 11
US-09-802-807-1/c
;; Sequence 1, Application US/09802807
;; Patent No. US20010034044A1
;; GENERAL INFORMATION:
;; APPLICANT: Treco, Douglas A.
;; APPLICANT: Hearle, Michael W.
;; APPLICANT: Seiden, Richard F.
;; TITLE OF INVENTION: GENOMIC SEQUENCES FOR PROTEIN PRODUCTION AND DELIVERY
;; FILE REFERENCE: 07236/016001
;; CURRENT APPLICATION NUMBER: US/09/802,807
;; CURRENT FILING DATE: 2001-03-08
;; PRIOR APPLICATION NUMBER: 60/084,663
;; PRIOR FILING DATE: 1998-05-07
;; NUMBER OF SEQ ID NOS: 7
;; SOFTWARE: FastSeq for Windows Version 3.0
;; SEQ ID NO 1
;; LENGTH: 7622
;; TYPE: DNA
;; ORGANISM: Homo sapiens
US-09-802-807-1

Query Match 26.7%; Score 26.4; DB 10; Length 7622;
Best Local Similarity 65.0%; Pred. No. 5.8;
Matches 39; Conservative 0; Mismatches 21; Indels 0; Gaps 0;

OY 19 GCATGTTGGCATGCTGCTAAGACCACTAGATCCAGACGCCAGCCCTTCTTT 78
DB 174 GCATGTTGGCGCAGCTGCTCTCAAAATCCTAGACTCAAGCAATCCACCTTGGCTT 115

RESULT 12
US-09-954-531-665
;; Sequence 665, Application US/09954531
;; Patent No. US20020165180A1
;; GENERAL INFORMATION:
;; APPLICANT: Weaver, Zoe
;; TITLE OF INVENTION: Process for Identifying Anti-Cancer Therapeutic Agents Using c
;; FILE REFERENCE: 689290-77
;; CURRENT APPLICATION NUMBER: US/09/954,531
;; CURRENT FILING DATE: 2002-05-02
;; PRIOR APPLICATION NUMBER: US/60/233,133
;; PRIOR FILING DATE: 2000-09-18
;; PRIOR APPLICATION NUMBER: US/60/234,009
;; PRIOR FILING DATE: 2000-09-20
;; PRIOR APPLICATION NUMBER: US/60/234,034
;; PRIOR FILING DATE: 2000-09-20
;; PRIOR APPLICATION NUMBER: US/60/234,509
;; PRIOR FILING DATE: 2000-09-22
;; PRIOR APPLICATION NUMBER: US/60/234,567


```

? PRR FILING DATE: 2000-09-22
? NUMBER OF SEQ ID NOS: 1392
? SOFTWARE: PatentIn version 3.0
? SEQ ID NO: 665
? LENGTH: 424
? TYPE: DNA
? ORGANISM: Homo sapiens
? FEATURE:
? NAME/KEY: misc.feature
? LOCATION: (1)..(424)
? OTHER INFORMATION: n=a,t,g or c
US-09-954-531-665

```

Query Match	26.5%	Score	26.2	DB	9	Length	424
Best Local Similarity	58.2%	Pred. No.	3				
Matches	46	Conservative	0	Mismatches	33	Indels	0
						Gaps	0

QY	17	AGCAATGTTGGGCATCTGCTAAAGGACCACTGATCCAGCAGGCCAGCCCTTCT	76
Db	288	AGCAAGGCGCTGAGCAAACTCTGCTACCAACAGAACCACTCTCTCAGCAGCTTACA	347
QY	77	TTCGTCTCAGAAAGCGCG	95
Db	348	TTCGAGTCAATACAGAGGG	366

RESULT 13
US-09-764-891-8344/C
; Sequence 8344, Application US/09764891
; Publication No. US20030077808A1
ORIGINAL INFORMATION

```

?
? TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
?
? FILE REFERENCE: PC006
?
? CURRENT APPLICATION NUMBER: US/09/764,891
?
? PRIORITY FILING DATE: 2001-01-17
? Prior application data removed - consult PALM or file wrapper
?
? NUMBER OF SEQ ID NOS: 1031
?
? SOFTWARE: PatentIn Ver. 2.0
?
? SEQ ID NO 8344
?
? LENGTH: 14759
?
? TYPE: DNA
?
? ORGANISM: Homo sapiens
?
? US-09-764-891-8344

```

Query Match	26.5%	Score 26.2;	DB 9;	Length 14759;
Best Local Similarity	67.3%	Pred. No. 8.4;		
Matches 37; Conservative	0;	Mismatches 18;	Indels 0;	Gaps 0;

Qy 20 CAAGTGTGGCATGCTGTAAAGCACCACATGAATCCAGCAGGCCAGACCCTTT 74
| | | | | | | | | | | | | | | | | |
Db 1216 CCATCTTGCCAGCGCTGCTCTTAAGACTCCCTAGACTCACGCATTCACGCACTTT 1162

RESULT 14
US-10-091-572-740/c
: Sequence 740, Application US/10091572
: Publication NO. US20030054373A1
: GENERAL INFORMATION:
: APPLICANT: Rosen et al.
: TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
: FILE REFERENCE: P1118C1
: CURRENT APPLICATION NUMBER: US/10/091,572
: CURRENT FILING DATE: 2002-03-07
: PRIOR APPLICATION NUMBER: 09/764,850
: PRIOR FILING DATE: 2001-01-17
: PRIOR APPLICATION NUMBER: 60/179,065
: PRIOR FILING DATE: 2000-01-31
: PRIOR APPLICATION NUMBER: 60/180,628
: PRIOR FILING DATE: 2000-02-04
: PRIOR APPLICATION NUMBER: 60/214,886
: PRIOR FILING DATE: 2000-06-28
: PRIOR APPLICATION NUMBER: 60/217,487

PRIOR FILING DATE:	2000-07-11
PRIOR APPLICATION NUMBER:	60/225,758
PRIOR FILING DATE:	2000-08-14
PRIOR APPLICATION NUMBER:	60/220,963
PRIOR FILING DATE:	2000-07-26
PRIOR APPLICATION NUMBER:	60/217,496
PRIOR FILING DATE:	2000-07-11
PRIOR APPLICATION NUMBER:	60/225,447
PRIOR FILING DATE:	2000-08-14
PRIOR APPLICATION NUMBER:	60/218,290
PRIOR FILING DATE:	2000-07-14
PRIOR APPLICATION NUMBER:	60/225,757
PRIOR FILING DATE:	2000-08-14
PRIOR APPLICATION NUMBER:	60/226,868
PRIOR FILING DATE:	2000-08-22
PRIOR APPLICATION NUMBER:	60/216,647
PRIOR FILING DATE:	2000-07-07
PRIOR APPLICATION NUMBER:	60/225,267
PRIOR FILING DATE:	2000-08-14
PRIOR APPLICATION NUMBER:	60/216,880
PRIOR FILING DATE:	2000-07-07
PRIOR APPLICATION NUMBER:	60/225,270
PRIOR FILING DATE:	2000-08-14
PRIOR APPLICATION NUMBER:	60/225,869
PRIOR FILING DATE:	2000-12-08
PRIOR APPLICATION NUMBER:	60/235,834
PRIOR FILING DATE:	2000-09-27
PRIOR APPLICATION NUMBER:	60/234,274
PRIOR FILING DATE:	2000-09-21
PRIOR APPLICATION NUMBER:	60/234,223
PRIOR FILING DATE:	2000-09-21
PRIOR APPLICATION NUMBER:	60/228,924
PRIOR FILING DATE:	2000-08-30
PRIOR APPLICATION NUMBER:	60/224,518
PRIOR FILING DATE:	2000-08-14
PRIOR APPLICATION NUMBER:	60/236,369
PRIOR FILING DATE:	2000-09-29
PRIOR APPLICATION NUMBER:	60/224,519
PRIOR FILING DATE:	2000-08-14
PRIOR APPLICATION NUMBER:	60/220,964
PRIOR FILING DATE:	2000-07-26
PRIOR APPLICATION NUMBER:	60/241,809
PRIOR FILING DATE:	2000-10-20
PRIOR APPLICATION NUMBER:	60/249,299
PRIOR FILING DATE:	2000-11-17
PRIOR APPLICATION NUMBER:	60/236,377
PRIOR FILING DATE:	2000-09-29
PRIOR APPLICATION NUMBER:	60/241,785
PRIOR FILING DATE:	2000-10-20
PRIOR APPLICATION NUMBER:	60/244,617
PRIOR FILING DATE:	2000-11-01
PRIOR APPLICATION NUMBER:	60/225,268
PRIOR FILING DATE:	2000-08-14
PRIOR APPLICATION NUMBER:	60/236,368
PRIOR FILING DATE:	2000-09-29
PRIOR APPLICATION NUMBER:	60/251,856
PRIOR FILING DATE:	2000-12-08
PRIOR APPLICATION NUMBER:	60/251,868
PRIOR FILING DATE:	2000-12-08
PRIOR APPLICATION NUMBER:	60/229,344
PRIOR FILING DATE:	2000-09-01
PRIOR APPLICATION NUMBER:	60/234,997
PRIOR FILING DATE:	2000-09-25
PRIOR APPLICATION NUMBER:	60/229,343
PRIOR FILING DATE:	2000-09-01
PRIOR APPLICATION NUMBER:	60/229,345
PRIOR FILING DATE:	2000-09-01
PRIOR APPLICATION NUMBER:	60/229,287
PRIOR FILING DATE:	2000-09-01
PRIOR APPLICATION NUMBER:	60/229,513
PRIOR FILING DATE:	2000-09-05
PRIOR APPLICATION NUMBER:	60/231,413
PRIOR FILING DATE:	2000-09-08

Query Match	26.5%;	Score 26.2;	DB 9;	Length 30420;
Best Local Similarity	54.7%;	Pred. No. 10;		
Matches 52;	Conservative 0;	Mismatches 43;	Indels 0;	Gaps 0;

Oy 4 CCCCTGAGCGCAAGGCAATGTTGGCATGCTGTAAGAGACCCTAGATCCAGACCC 63
|| ||| ||||| | | | ||| | ||| |||
Db 6659 CCACGTGAACAAAGGAGGAGGTTGAATCTGTTTAAAGAACATTTAGATTCACAGTTG 6600
Oy 64 CAGACCCCTTTCTTTGTTGTCAGAAAGGCGGAA 98
|| ||| || | | | | | |||
Db 6599 CAATCCCTGCTCTCTCTCTTCAGAACACAGGAA 6565

Search completed: May 2, 2003, 14:43:11
Job time : 94.8862 secs

GenCore version 5.1.5
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 2, 2003, 11:21:20 ; Search time 411.145 Seconds

(without alignments)
6795.341 Million cell updates/sec

Title: US-09-719-554-3_COPY_3065_3160

Perfect score: 96
Sequence: 1 gccgcctggcactccttgagg.....tgagtgagtgagtcataagt 96

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl: *
1: gb_ba: *
2: gb_htg: *
3: gb_in: *
4: gb_om: *
5: gb_ov: *
6: gb_pat: *
7: gb_ph: *
8: gb_pl: *
9: gb_pr: *
10: gb_ro: *
11: gb_sts: *
12: gb_sy: *
13: gb_un: *
14: gb_vl: *
15: em_ba: *
16: em_fun: *
17: em_hum: *
18: em_in: *
19: em_mu: *
20: em_om: *
21: em_or: *
22: em_ov: *
23: em_pat: *
24: em_ph: *
25: em_pl: *
26: em_ro: *
27: em_sts: *
28: em_un: *
29: em_vl: *
30: em_htg_hum: *
31: em_htg_inv: *
32: em_htg_other: *
33: em_htg_mus: *
34: em_htg_pln: *
35: em_htg_rtd: *
36: em_htg_mam: *
37: em_htg_vrt: *
38: em_sy: *
39: em_htgo_hum: *
40: em_htgo_mus: *
41: em_htgo_other: *

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	96	100.0	1321	6	AX000956	AX000956 Sequence
2	96	100.0	1321	6	AX027470	AX027470 Sequence
3	96	100.0	1321	6	AF072500	AF072500 Homo sapi
4	96	100.0	1326	6	AX007979	AX007979 Sequence
5	96	100.0	2372	6	AX000965	AX000965 Sequence
6	96	100.0	2372	6	AX027479	AX027479 Sequence
7	96	100.0	2372	6	AF072504	AF072504 Homo sapi
8	96	100.0	10499	6	AX007980	AX007980 Sequence
9	96	100.0	56093	6	AX329572	AX329572 Sequence
10	96	100.0	56093	6	HSAC000064	AC000064 Human BAC
11	96	100.0	149194	9	AC007566	AC007566 Homo sapi
12	95.6	99.6	7582	6	AX027480	AX027480 Sequence
13	95.6	98.3	2838	6	AX000957	AX000957 Sequence
14	94.4	98.3	2838	6	AX027471	AX027471 Sequence
15	94.4	98.3	2838	6	AF072499	AF072499 Homo sapi
16	94.4	98.3	2838	6	AF072499	AF072499 Homo sapi
17	71.4	74.4	139744	9	AL133513	AL133513 Human DNA
18	70.8	73.8	104853	9	AC117444	AC117444 Homo sapi
19	70.8	73.8	169462	2	AC024033	AC024033 Homo sapi
20	70.8	73.8	186723	2	AC027752	AC027752 Homo sapi
21	70.4	73.3	1774	6	AF007994	AF007994 Sequence
22	70.4	73.3	40205	9	AF045450	AF045450 Homo sapi
23	70.4	73.3	142742	9	AF121782	AF121782 Homo sapi
24	70.4	73.3	251124	9	HUAE000660	AE000660 Homo sapi
25	70.4	73.3	340000	9	HS21C080	AL163280 Homo sapi
26	69.8	72.7	94585	9	HS0J401P4	AL109922 Human DNA
27	69.6	72.5	153803	2	AC055722	AC055722 Homo sapi
28	69.6	72.5	164310	2	AC016156	AC016156 Homo sapi
29	69.6	72.5	182224	2	AC093116	AC093116 Homo sapi
30	69.6	72.5	203230	2	AC026346	AC026346 Homo sapi
31	69.2	72.1	70352	9	AL592310	AL592310 Human DNA
32	69.2	72.1	25437	9	AF000500	AF000500 Homo sapi
33	69.2	72.1	270000	9	AB026898	AB026898 Homo sapi
34	68.8	71.7	66369	2	AC090325	AC090325 Homo sapi
35	68.8	71.7	101196	9	AC087436	AC087436 Homo sapi
36	68.8	71.7	136901	9	AC073626	AC073626 Homo sapi
37	68.8	71.7	147655	9	AC104716	AC104716 Homo sapi
38	68.8	71.7	150955	9	AL356111	AL356111 Human DNA
39	68.8	71.7	155925	9	AC096536	AC096536 Homo sapi
40	68.8	71.7	157043	2	AC013692	AC013692 Homo sapi
41	68.8	71.7	161571	2	AC092708	AC092708 Homo sapi
42	68.8	71.7	161938	9	AL360227	AL360227 Human DNA
43	68.8	71.7	166107	2	AC024427	AC024427 Homo sapi
44	68.8	71.7	166700	9	AC013759	AC013759 Homo sapi
45	68.8	71.7	169072	9	AC009946	AC009946 Homo sapi

ALIGNMENTS

RESULT 1
AX000956 1321 bp DNA linear PAT 10-MAR-2000
LOCUS AX000956
DEFINITION Sequence 1 from Patent WO9902656.
ACCESSION AX000956
VERSION AX000956.1 GI:7241198
KEYWORDS
SOURCE
ORGANISM
REFERENCE 1 (bases 1 to 1321)
AUTHORS Beseme, F. and Blond, J.
TITLE ENDOGENETIC RETROVIRAL SEQUENCES, ASSOCIATED WITH AUTOIMMUNE
DISEASES OR WITH PREGNANCY DISORDERS
JOURNAL Patent: WO 9902656-A 1 21-JAN-1999;

Pred. No. is the number of results predicted by chance to have a

FEATURES BIO MERIEUX (FR); BESEME FREDERIC (FR)
source Location/Qualifiers
1.1321
/organism="unidentified"
/db_xref="taxon:32644"
BASE COUNT 365 a 352 c 277 g 327 t
ORIGIN

Query Match 100.0%; Score 96; DB 6; Length 1321;
Best Local Similarity 100.0%; Pred. No. 7.9e-19;
Matches 96; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCGCTGGACCTCTGAGGAGATATAATTAACACCATCTTACAGCTAGACCTCTT 60
DB 717 GCCGCTGGACCTCTGAGGAGATATAATTAACACCATCTTACAGCTAGACCTCTT 776

QY 61 TTGTAGAAAAGCAATGAGTGAAGTCCCTAAGT 96
DB 777 TTGTAGAAAAGCAATGAGTGAAGTCCCTAAGT 812

RESULT 2
AX027470 1321 bp DNA linear PAT 16-SEP-2000
LOCUS Sequence 20 from Patent FR2788784.
DEFINITION AX027470
ACCESSION AX027470.1 GI:10188434
VERSION
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 1321)
AUTHORS Mallet,F., Voisset,C. and Paranhos,B.G.
JOURNAL Patent: FR 2788784-A 20 28-JUL-2000;
BIO MERIEUX (FR)

FEATURES
source Location/Qualifiers
1.1321
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 365 a 352 c 277 g 327 t
ORIGIN

Query Match 100.0%; Score 96; DB 6; Length 1321;
Best Local Similarity 100.0%; Pred. No. 7.9e-19;
Matches 96; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCGCTGGACCTCTGAGGAGATATAATTAACACCATCTTACAGCTAGACCTCTT 60
DB 717 GCCGCTGGACCTCTGAGGAGATATAATTAACACCATCTTACAGCTAGACCTCTT 776

QY 61 TTGTAGAAAAGCAATGAGTGAAGTCCCTAAGT 96
DB 777 TTGTAGAAAAGCAATGAGTGAAGTCCCTAAGT 812

RESULT 3
AF072500 1321 bp mRNA linear PRI 10-FEB-1999
LOCUS Homo sapiens endogenous retrovirus W sequence.
DEFINITION AF072500
ACCESSION AF072500.1 GI:4262284
VERSION
KEYWORDS
SOURCE Homo sapiens.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 1321)
AUTHORS Blond,J.L., Beseme,F., Durel,L., Bouton,O., Bedin,F., Perron,H.,
Mandrand,B. and Mallet,F.
JOURNAL Molecular characterization and placental expression of HERV-W, a
new human endogenous retrovirus family
J. VIROL. 73 (2), 1175-1185 (1999)

MEDLINE 99099005
PUBMED 9882319
REFERENCE 2 (bases 1 to 1321)
AUTHORS Blond,J.L., Beseme,F. and Mallet,F.
TITLE Direct Submission
JOURNAL Submitted (19-JUN-1998) UM103 CNRS-Biomerieux, ENS Lyon, 46 allée
d'Italie, Lyon, Cedex 07 69364, France

FEATURES
source Location/Qualifiers
1.1321
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="cl.6a2"
/tissue_type="placenta"
/clone_lib="Clontech 5'-stretch plus library, Cat number
HLS014a"
misc_feature 1.1321
/note="corresponds to gag"
repeat_region 1.1321
/rpt_family="HERV-W"
5'UTR 1.120
/note="5' end of R uncertain"
5'UTR 121..574
/note="05"
578..595
primer_bind /note="putative primer binding site for tRNA-W"
BASE COUNT 365 a 352 c 277 g 327 t
ORIGIN

Query Match 100.0%; Score 96; DB 9; Length 1321;
Best Local Similarity 100.0%; Pred. No. 7.9e-19;
Matches 96; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCGCTGGACCTCTGAGGAGATATAATTAACACCATCTTACAGCTAGACCTCTT 60
DB 717 GCCGCTGGACCTCTGAGGAGATATAATTAACACCATCTTACAGCTAGACCTCTT 776

QY 61 TTGTAGAAAAGCAATGAGTGAAGTCCCTAAGT 96
DB 777 TTGTAGAAAAGCAATGAGTGAAGTCCCTAAGT 812

RESULT 4
AX007979 1326 bp DNA linear PAT 06-SEP-2000
LOCUS Sequence 2 from Patent WO967395.
DEFINITION AX007979
ACCESSION AX007979
VERSION AX007979.1 GI:9995676
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 1326)
AUTHORS Perrin,J.P., Rieger,F. and Alliel,P.M.
TITLE Nucleic sequence and deduced protein sequence family with human
endogenous retroviral motifs, and their uses
JOURNAL INST NAT SANTE RECH MED (FR); PERIN JEAN PIERRE (FR); RIEGER
FRANCOIS (FR); ALLIEL PATRICK M (FR)

FEATURES
source Location/Qualifiers
1.1326
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 425 a 306 c 303 g 292 t
ORIGIN

Query Match 100.0%; Score 96; DB 6; Length 1326;
Best Local Similarity 100.0%; Pred. No. 7.9e-19;
Matches 96; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCGCTGGACCTCTGAGGAGATATAATTAACACCATCTTACAGCTAGACCTCTT 60
DB 1 GCCGCTGGACCTCTGAGGAGATATAATTAACACCATCTTACAGCTAGACCTCTT 60


```

REFERENCE 1 (bases 1 to 10499)
AUTHORS Perlin,J.P., Rieger,F. and Alliel,P.M.
TITLE Nucleic sequence and deduced protein sequence family with human
JOURNAL INST NAT SANTE RECH MED (FR); PERIN JEAN PIERRE (FR); RIEGER
FRANCOIS (FR); ALLIEL PATRICK M (FR)
FEATURES
    source
        1. 10499
            /db_xref="taxon:9606"
BASE COUNT 3048 a 2676 c 2280 g 2495 t
ORIGIN
Query Match 100.0%; Score 96; DB 6; Length 10499;
Best Local Similarity 100.0%; Pred. No. 5.5e-19;
Matches 96; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GCCGCGCTGGACATCTCTGAGGAGATTAATTAATTAACACCATCTTACAGCTAGCTCTT 60
    |||||||
Db 3065 GCCGCGCTGGACATCTCTGAGGAGATTAATTAATTAACACCATCTTACAGCTAGCTCTT 3124
    |||||||
QY 61 TTGTAGAAAAGGCAAAATGAGTGAAGTGCATTAAGT 96
    |||||||
Db 3125 TTGTAGAAAAGGCAAAATGAGTGAAGTGCATTAAGT 3160
    |||||||
RESULT 9
AX329572 56093 bp DNA linear PAT 09-JAN-2002
LOCUS AX329572
DEFINITION Sequence 81 from Patent WO0194629.
ACCESSION AX329572
VERSION AX329572.1 GI:18102550
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1
    Young,P.E., Augustus,M., Carter,K.C., Ebner,R., Endress,G.,
    Horigan,S., Soppet,D.R. and Weaver,Z.
    Cancer gene determination and therapeutic screening using signature
    gene sets
    Patent: WO 0194629-A 81 13-DEC-2001;
JOURNAL Avalon Pharmaceuticals (US)
FEATURES
    source
        1. 56093
            /organism="Homo sapiens"
            /db_xref="taxon:9606"
BASE COUNT 16164 a 12346 c 10702 g 16881 t
ORIGIN
Query Match 100.0%; Score 96; DB 6; Length 56093;
Best Local Similarity 100.0%; Pred. No. 4.1e-19;
Matches 96; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GCCGCGCTGGACATCTCTGAGGAGATTAATTAATTAACACCATCTTACAGCTAGCTCTT 60
    |||||||
Db 31065 GCCGCGCTGGACATCTCTGAGGAGATTAATTAATTAACACCATCTTACAGCTAGCTCTT 31124
    |||||||
QY 61 TTGTAGAAAAGGCAAAATGAGTGAAGTGCATTAAGT 96
    |||||||
Db 31125 TTGTAGAAAAGGCAAAATGAGTGAAGTGCATTAAGT 31160
    |||||||
RESULT 10
HSAC000064 56093 bp DNA linear PRI 13-NOV-1996
LOCUS HSAC000064
DEFINITION Human BAC clone RG083M05 from 7q21-7q22, complete sequence.
ACCESSION AC000064
VERSION AC000064.1 GI:1669369
KEYWORDS HTG.
SOURCE Homo sapiens.

```

```

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 56093)
AUTHORS Pauley,A.
TITLE The sequence of H. sapiens BAC clone RG083M05
JOURNAL Unpublished (1996)
REFERENCE 2 (bases 1 to 56093)
AUTHORS Waterston,R.
TITLE Direct Submission
JOURNAL Submitted (13-NOV-1996)
COMMENT Genome Sequencing Center
    Department of Genetics, Washington University
    St. Louis, MO 63108, USA
    e-mail: sapiens@watson.wustl.edu

NOTICE: This sequence may not represent the entire insert of this
clone. It may be shorter because we only sequence overlapping
sections once, or longer because we provide a small overlap between
neighboring submissions.

This sequence was finished as follows unless otherwise noted:
all regions were double stranded or sequenced with an alternate
chemistry; an attempt was made to resolve all sequencing problems,
such as compressions and repeats; all regions were covered by
sequence from more than one subclone; and the assembly was
confirmed by restriction digest.

SOURCE INFORMATION:
This clone is from the first release of the human BAC library. The
library contains cloned DNA from a human male fibroblast cell line
978SK. For references see: Shizuya et al., Proc. Natl. Acad. Sci.
89:8794-8797 (1992); Kim et al., Genomics 34:213-218 (1996).
VECTOR: pBelo
Selection: chloramphenicol

NEIGHBORING SEQUENCE INFORMATION:
The orientation of this clone is unknown. Actual start of this
clone is at base position 1 of H_RG083M05, actual end is at 56093
of H_RG083M05

This clone contains STS SW551725.
FEATURES
    source
        1. 56093
            /organism="Homo sapiens"
            /db_xref="taxon:9606"
            /chromosome="7"
            /map="7q21-7q22"
            /clone="H_RG083M05"
            /clone_lib="CITB-978SK-B"
            complement(838..1131)
            <1360..16971
            /rpt_family="ALU"
            /gene="WUGSC:H_RG083M05.1"
            join(<1360..1503,4181..4370,4587..4774,6422..6556,
            9483..9547,11631..11773,11864..12021,13131..13296,
            14885..14988,16349..16546,16837..16971)
            /gene="WUGSC:H_RG083M05.1"
            /note="ATPase, strong similarity to peroxisome
            biosynthesis protein PAB1 (PID:q1172019); coded for by
            human cDNA C04279 (NID:q1467530)"
            /codon_start=1
            /protein_id="AAB46346.1"
            /db_xref="GI:1669371"
            /translation="KRLENIQKLEVAESEAVMOPSVVLDDDLINGLPAVPEHEH
            SPDAERCEIILNVINKLDCDINKFTDLDLHVAEETGFAVDFTVLDRAISHRL
            SROSISTREKLVLTLDFOKALRGFLPALSRSVNLHPRDLQMDKIGGHEVROILMD
            TIOLEPAKVLKREKYEPELPANLPIRQRTGILYGPQGTLLAGVIASESMNFISY
            KQBELSKYIGASEQAVRDIFIRAOAKRCILFPEFEISAPRGHDNGVDRVNO
            LKQLDGVGICGVYLAATSRPDLIPALLRPRGRDKVVCPPPOVYVSYLESKTO
            QMHSLVSRLEILNVLSDSLPLADDVDLDQHYASTDSTGTGDLKALYNAQLEALHG
            MLTKSKSEILPDESKFNMYRLYFGSSYESELNGVSSDLSOCLAPSSMTDLPQVP
            GKDOLFSQPPVLRTASQEGCELTQEOBQLRADISIIIGRYRSOSGEDESNMGPI

```


KTRIAISQSHLMTALGHTRPISSEDWKNFAEL"

repeat_region complement(4948. .5130)
/rpt_family="ALU"

repeat_region complement(6581. .7133)
/rpt_family="L1"

repeat_region complement(7767. .8037)
/rpt_family="ALU"

repeat_region complement(8186. .8472)
/rpt_family="ALU"

misc_feature 8473. .8625
/gene="WUGSC:H.RG083M05.1"

misc_feature 287-444"
/note="match to human 3' EST H75782 (NID:g1049794), bases 287-444"

misc_feature 8841. .9161
/gene="WUGSC:H.RG083M05.1"

misc_feature 21-348"
/note="match to human 5' EST H75921 (NID:g1050050), bases 21-348"

misc_feature 9481. .9547
/gene="WUGSC:H.RG083M05.1"

repeat_region 276-343"
/note="match to human 5' EST N22627 (NID:g1130501), bases 276-343"

misc_feature complement(12612. .12907)
/rpt_family="ALU"

misc_feature 13670. .13793
/gene="WUGSC:H.RG083M05.1"

repeat_region 143-266"
/note="match to human 5' EST H41382 (NID:917434), bases 143-266"

repeat_region 13794. .13877
/rpt_family="ALU"

misc_feature 13878. .13906
/gene="WUGSC:H.RG083M05.1"

repeat_region 30-58"
/note="match to human 5' EST H41382 (NID:g117434), bases 30-58"

repeat_region 13907. .14104
/rpt_family="ALU"

repeat_region complement(14110. .14137)
/rpt_family="L1"

repeat_region complement(15618. .15907)
/rpt_family="ALU"

repeat_region 17227. .17522
/rpt_family="ALU"

misc_feature 18667. .19235
/note="match to human fetal brain 5' EST D61494 (NID:g970409), bases 1-255, and to human 3' EST R07476 (NID:g759399)"

repeat_region 19550. .19670
/rpt_family="ALU"

misc_feature 21507. .37303
/note="similarity to various SS-RNA virus polypeptides; pseudogene: region of matches and close matches to multiple human ESTs, see R68740 (NID:g842257)"

exon 37316. .37489
/note="Grail prediction, score = 80"

repeat_region /evidence=not_experimental
complement(38938. .39224)
/rpt_family="ALU"

misc_feature 39325. .39707
/note="match to multiple human ESTs, see N30113 (NID:g1148633)"

repeat_region 39800. .40085
/rpt_family="ALU"

repeat_region complement(40247. .40538)
/rpt_family="ALU"

repeat_region complement(40924. .40924)
/rpt_family="ALU"

repeat_region complement(42283. .42891)
/rpt_family="ALU"

repeat_region complement(45474. .45613)
/rpt_family="ALU"

misc_feature complement(45614. .45737)
/note="match to human 3' EST H48898 (NID:g988738), bases 129-333"

misc_feature complement(46107. .47026)

/note="match to multiple human ESTs, see N81064 (NID:g1243765), H48897 (NID:g988737), and M78831 (NID:g273146)"

repeat_region complement(47027. .47318)
/rpt_family="ALU"

misc_feature complement(47365. .47782)
/note="match to multiple human ESTs, see M37495 (NID:g1318089)"

misc_feature 47898. .48115
/note="match to human 5' EST H62306 (NID:g1015138), bases 93-368"

repeat_region complement(48116. .48405)
/rpt_family="ALU"

misc_feature complement(48406. .48584)
/note="match to human 3' EST N29952 (NID:g1148472), bases 290-455, and 5' EST R12730 (NID:g765806)"

repeat_region complement(48787. .49405)
/rpt_family="ALU"

misc_feature complement(49406. .49534)
/note="match to human 3' EST R65794 (NID:g838432), bases 309-440"

repeat_region complement(49638. .49672)
/rpt_family="ALU"

misc_feature complement(49674. .49890)
/note="match to human 3' EST N29952 (NID:g1148472) and 5' EST N29938 (NID:g1148458), sequences are from opposite ends of the same clone"

gene complement(49698. .51806)
/gene="WUGSC:H.RG083M05.2"

CDS complement(join(49698. .49888, 51575. .51806))
/gene="WUGSC:H.RG083M05.2"

exon /note="coded for by human cDNAs M7389 (NID:g1139205), R65891 (NID:g838529), R65794 (NID:g838432) and R65794 (NID:g838432)"

repeat_region /codon_start=1
/protein_id="AAB46345.1"

misc_feature /db_xref="GI:1669370"

repeat_region /translation="MEFYRQCGIIFRPGVYVQIDVYSVIDEGRKYAQRGF
IIDDYCKRSALTLWLTPLSSPRDQFPASTIIQPEDLPKMYELFVCHAPSEYK
SRSSPFPTPTPEKGYITWTHVGEPPIITIKESVANH"

exon complement(51576. .51758)
/gene="WUGSC:H.RG083M05.2"

repeat_region /note="Grail prediction, score = 86"

misc_feature /evidence=not_experimental
complement(52052. .52329)
/rpt_family="L1"

misc_feature 55557. .55843
/note="match to human EST M79192 (NID:g273505) base 2-289"

Query Match 100.0%; Score 96; DB 9; Length 56093;
Best Local Similarity 100.0%; Pred. No. 4,1e-19;
Matches 96; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCGCTGCGACTCCTCGAGGAGATATAATTATACACATCTTACAGTACACTCTT 60
|||||
Db 31065 GCCGCTGCGACTCCTCGAGGAGATATAATTATACACATCTTACAGTACACTCTT 31124
|||||

QY 61 TTGTAGAAAGCAAAATGAGTGAAGTCCATAACT 96
|||||
Db 31125 TTGTAGAAAGCAAAATGAGTGAAGTCCATAACT 31160
|||||

RESULT 11
AC007566/c 149194 bp DNA linear PRI 01-MAR-2002
LOCUS Homo sapiens BAC clone CTB-1065 from 7q21-7q22, complete sequence.
DEFINITION AC007566
ACCESSION AC007566.2 GI:11181861
VERSION
KEYWORDS HIG.
SOURCE Homo sapiens.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 149194)
 AUTHORS Sulston,J.E. and Waterston,R.
 TITLE Toward a complete human genome sequence
 JOURNAL Genome Res. 8 (11), 1097-1108 (1998)
 MEDLINE 99063792
 PUBMED 9847074
 REFERENCE 2 (bases 1 to 149194)
 AUTHORS Du,Z.
 TITLE The sequence of Homo sapiens BAC clone CTB-10G5
 JOURNAL Unpublished (2001)
 REFERENCE 3 (bases 1 to 149194)
 AUTHORS Waterston,R.H.
 TITLE Direct Submission
 JOURNAL Submitted (15-MAY-1999) Genome Sequencing Center, Washington University School of Medicine, 4444 Forest Park Parkway, St. Louis, MO 63108, USA
 REFERENCE 4 (bases 1 to 149194)
 AUTHORS Waterston,R.
 TITLE Direct Submission
 JOURNAL Submitted (02-OCT-2000) Department of Genetics, Washington University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
 REFERENCE 5 (bases 1 to 149194)
 AUTHORS Waterston,R.H.
 TITLE Direct Submission
 JOURNAL Submitted (16-NOV-2000) Genome Sequencing Center, Washington University School of Medicine, 4444 Forest Park Parkway, St. Louis, MO 63108, USA
 REFERENCE 6 (bases 1 to 149194)
 AUTHORS Waterston,R.H.
 TITLE Direct Submission
 JOURNAL Submitted (03-JAN-2002) Genome Sequencing Center, Washington University School of Medicine, 4444 Forest Park Parkway, St. Louis, MO 63108, USA
 REFERENCE 7 (bases 1 to 149194)
 AUTHORS Waterston,R.H.
 TITLE Direct Submission
 JOURNAL Submitted (06-FEB-2002) Genome Sequencing Center, Washington University School of Medicine, 4444 Forest Park Parkway, St. Louis, MO 63108, USA
 REFERENCE 8 (bases 1 to 149194)
 AUTHORS Waterston,R.
 TITLE Direct Submission
 JOURNAL Submitted (01-MAR-2002) Department of Genetics, Washington University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
 On Nov 16, 2000 this sequence version replaced gi:4835815.
 ----- Genome Center
 Center: Washington University Genome Sequencing Center
 Center code: WUGSC
 Web site: <http://genome.wustl.edu/gsc>
 Contact: saplens@wustl.wustl.edu
 ----- Summary Statistics
 Center project name: H_RG010G05

NOTICE: This sequence may not represent the entire insert of this clone. It may be shorter because we only sequence overlapping clone sections once, or longer because we provide a small overlap between neighboring data submissions.

This sequence was finished as follows unless otherwise noted:
 all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

MAPPING INFORMATION:

The sequence of this clone was established as part of a mapping and sequencing collaboration between the NHGRI Chromosome 7 Mapping Project (Eric D. Green, Director), John D. McPherson in the Department of Genetics (Washington University), and the Washington University Genome Sequencing Center. For additional information

about the map position of this sequence, see
<http://www.nhgri.nih.gov/DIR/CTB/CHR7>, send
<mailto:egreen@nhgri.nih.gov>, or see <http://genome.wustl.edu/gsc>

SOURCE INFORMATION:

Clone CTB-10G5 is from the first release of the human BAC library CTB-978SK-B. The library contains cloned DNA from the male fibroblast cell line 978SK. See: Shizuya et al., Proc. Natl. Sci. USA 89:8794-7 (1992); U-J. Kim et al., Genomics 34:213-8 (1996). This clone is available from Research Genetics, Inc. (<http://www.resgen.com>).
 VECTOR: pRelBAC11
 Selection: chloramphenicol

NEIGHBORING SEQUENCE INFORMATION:

The clone sequenced to the left is RP5-911H5, 200 base pair overlap. Actual start of this clone is at base position 195 of CTB-10G5; actual end is at base position 150532 of CTB-10G5.

FEATURES

source

The clone CTB-10G5 contains the entire sequence of CTB-83M5.
 Location/Qualifiers

```

1..149194
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="7"
/map="7q21-7q22"
/clone="CTB-10G5"
/clone_11b="CTB-978SK-B"

1..1634
/rpt_family="L1"
misc_feature
2248..2388
/note="match to EST BG752883 (NID:g14063536)"
misc_feature
2248..2387
/note="match to EST AA149693 (NID:g1720635) zn99d06.r1"
misc_feature
2248..2387
/note="match to EST AM579261 (NID:g7254310)"
misc_feature
2248..2387
/note="match to EST BG766882 (NID:g14077535)"
misc_feature
2248..2387
/note="match to EST B1160365 (NID:g14620366)"
misc_feature
2248..2374
/note="similar to Homo sapiens EST BF758865 (NID:g12106765)"
misc_feature
2248..2287
/note="match to EST AV686676 (NID:g10288539)"
misc_feature
2253..2387
/note="match to EST AU123510 (NID:g10948226)"
misc_feature
2344..2387
/note="match to EST AV686676 (NID:g10288539)"
misc_feature
2696..3066
/rpt_family="MALR"
repeat_region
3108..3392
/rpt_family="ALU"
misc_feature
3540..3628
/note="match to EST AA425526 (NID:g2106267) zw48b03.r1"
misc_feature
3540..3628
/note="similar to Mus musculus EST BB253526 (NID:g8946272)"
misc_feature
3542..3632
/note="match to EST AV686676 (NID:g10288539)"
misc_feature
3542..3628
/note="match to EST AA149693 (NID:g1720635) zn99d06.r1"
misc_feature
3542..3628
/note="match to EST AU123510 (NID:g10948226)"
misc_feature
3542..3628
/note="match to EST BG766882 (NID:g14077535)"
misc_feature
3542..3628
/note="match to EST B1160365 (NID:g14620366)"
misc_feature
3542..3609
/note="match to EST AM579261 (NID:g7254310)"
misc_feature
3562..3628
/note="match to EST BE272564 (NID:g9146913)"
misc_feature
3598..3628
```

misc_feature /note="match to EST BE299708 (NID:g9183456)"
3714..3785
/note="match to EST BG260659 (NID:g12770475)"
3717..3785
misc_feature /note="similar to Mus musculus EST BE994936
(NID:g10678674)"
3719..3785
misc_feature /note="match to EST AA149693 (NID:g1206635) zn99d06.r1"
3719..3785
misc_feature /note="match to EST AA425526 (NID:g2106267) zw48b03.r1"
3719..3785
misc_feature /note="match to EST AU123510 (NID:g10948226)"
3719..3785
misc_feature /note="match to EST BE272564 (NID:g9146913)"
3719..3785
misc_feature /note="match to EST BE299708 (NID:g9183456)"
3719..3785
misc_feature /note="match to EST BG766882 (NID:g14077535)"
3719..3785
misc_feature /note="match to EST B1160365 (NID:g14620366)"
3719..3785
misc_feature /note="similar to Mus musculus EST BE253526
(NID:g8946272)"
misc_feature /note="match to EST AV686676 (NID:g10288539)"
4730..4827
repeat_region /rpt_family="L2"
4834..4925
repeat_region /rpt_family="L1"
5715..5860
misc_feature /note="match to EST AA425526 (NID:g2106267) zw48b03.r1"
5715..5860
misc_feature /note="match to EST AU123510 (NID:g10948226)"
5715..5860
misc_feature /note="match to EST BE272564 (NID:g9146913)"
5715..5860
misc_feature /note="match to EST BE299708 (NID:g9183456)"
5715..5860
misc_feature /note="match to EST BG260659 (NID:g12770475)"
5715..5860
misc_feature /note="match to EST BG766882 (NID:g14077535)"
5715..5860
misc_feature /note="match to EST B1160365 (NID:g14620366)"
5715..5860
misc_feature /note="similar to Mus musculus EST BE94936
100.0%; Score 96; DB 9; Length 149194;
Best Local Similarity 100.0%; Pred. No. 3.4e-19;
Matches 96; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GCCGCTGGCAGCTCTGAGGAGATATAATATATACACCATCTTACAGCTGACCTCTT 60
|||||
Db 90939 GCCGCTGGCAGCTCTGAGGAGATATAATATATACACCATCTTACAGCTGACCTCTT 90880
|||||
QY 61 TTGTAGAAAAGCAATGAGTGAAGTGCATTAAGT 96
|||||
Db 90879 TTGTAGAAAAGCAATGAGTGAAGTGCATTAAGT 90844
|||||
RESULT 12
AX000966 7582 bp DNA Linear PAT 10-MAR-2000
LOCUS AX000966
DEFINITION Sequence 11 from Patent WO9902696.
ACCESSION AX000966
VERSION AX000966.1 GI:7241208
KEYWORDS
SOURCE
ORGANISM
unidentified.
unclassified.
REFERENCE
1 (bases 1 to 7582)
AUTHORS
Beseme,F. and Blond,J.
TITLE
ENDOGENETIC RETROVIRAL SEQUENCES, ASSOCIATED WITH AUTOIMMUNE
DISEASES OR WITH PREGNANCY DISORDERS

JOURNAL Patent: WO 9902696-A 11 21-JAN-1999:
BIO MERIEUX (FR); BESEME FREDERIC (FR)
FEATURES
source
1..7582
/organism="unidentified"
/db_xref="taxon:32644"
BASE COUNT 2156 a 1876 c 1538 g 1796 t 216 others
ORIGIN
Query Match 99.6%; Score 95.6; DB 6; Length 7582;
Best Local Similarity 99.0%; Pred. No. 7.7e-19;
Matches 95; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 GCCGCTGGCAGCTCTGAGGAGATATAATATATACACCATCTTACAGCTGACCTCTT 60
|||||
Db 718 GCCGCTGGCAGCTCTGAGGAGATATAATATATACACCATCTTACAGCTGACCTCTT 777
|||||
QY 61 TTGTAGAAAAGCAATGAGTGAAGTGCATTAAGT 96
|||||
Db 778 TTGTAGAAAAGCAATGAGTGAAGTGCATTAAGT 813
|||||
RESULT 13
AX027480 7582 bp DNA Linear PAT 16-SEP-2000
LOCUS AX027480
DEFINITION Sequence 30 from Patent FR2788784.
ACCESSION AX027480
VERSION AX027480.1 GI:10188444
KEYWORDS
SOURCE
human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1 (bases 1 to 7582)
AUTHORS
Mallet,F., Voisset,C. and Paranhos,B.G.
JOURNAL
Patent: FR 2788784-A 30 28-JUL-2000;
BIO MERIEUX (FR)
FEATURES
source
1..7582
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 2156 a 1876 c 1538 g 1796 t 216 others
ORIGIN
Query Match 99.6%; Score 95.6; DB 6; Length 7582;
Best Local Similarity 99.0%; Pred. No. 7.7e-19;
Matches 95; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 GCCGCTGGCAGCTCTGAGGAGATATAATATATACACCATCTTACAGCTGACCTCTT 60
|||||
Db 718 GCCGCTGGCAGCTCTGAGGAGATATAATATATACACCATCTTACAGCTGACCTCTT 777
|||||
QY 61 TTGTAGAAAAGCAATGAGTGAAGTGCATTAAGT 96
|||||
Db 778 TTGTAGAAAAGCAATGAGTGAAGTGCATTAAGT 813
|||||
RESULT 14
AX000957 2938 bp DNA Linear PAT 10-MAR-2000
LOCUS AX000957
DEFINITION Sequence 2 from Patent WO9902696.
ACCESSION AX000957
VERSION AX000957.1 GI:7241199
KEYWORDS
SOURCE
ORGANISM
unidentified.
unclassified.
REFERENCE
1 (bases 1 to 2938)
AUTHORS
Beseme,F. and Blond,J.
TITLE
ENDOGENETIC RETROVIRAL SEQUENCES, ASSOCIATED WITH AUTOIMMUNE
DISEASES OR WITH PREGNANCY DISORDERS
Patent: WO 9902696-A 2 21-JAN-1999;
BIO MERIEUX (FR); BESEME FREDERIC (FR)

FEATURES Location/Qualifiers
Source 1..2938
/organism="unidentified"
/db_xref="taxon:32644"

BASE COUNT 878 a 720 c 646 g 692 t 2 others

ORIGIN

Query Match 98.3%; Score 94.4; DB 6; Length 2938;
Best Local Similarity 99.0%; Pred. No. 2.1e-18;
Matches 95; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GCCGCTGGCAGCTCTGAGGAGTATATATATACACCATCTTACAGCTAGACCTCTT 60
|||||
Db 133 GCCGCTGGCAGCTCTGAGGAGTATATATATACACCATCTTACAGCTAGACCTCTT 192
|||||

QY 61 TTGTAGAAAAGCAATGAGTGAAGTGCCTAAGT 96
|||||
Db 193 TTGTAGAAAAGCAATGAGTGAAGTGCCTAAGT 228
|||||

RESULT 15
AX027471 2938 bp DNA linear PAT 16-SEP-2000
LOCUS Sequence 21 from Patent FR2788784.
DEFINITION AX027471
ACCESSION AX027471
VERSION AX027471.1 GI:10188435
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 2938)
AUTHORS Mallet, F., Voisset, C. and Paranhos, B. G.
JOURNAL Patent: FR 2788784-A 21 28-JUL-2000;
BIO MERIEUX (FR)

FEATURES Location/Qualifiers
Source 1..2938
/organism="Homo sapiens"
/db_xref="taxon:9606"

BASE COUNT 878 a 720 c 646 g 692 t 2 others

ORIGIN

Query Match 98.3%; Score 94.4; DB 6; Length 2938;
Best Local Similarity 99.0%; Pred. No. 2.1e-18;
Matches 95; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GCCGCTGGCAGCTCTGAGGAGTATATATATACACCATCTTACAGCTAGACCTCTT 60
|||||
Db 133 GCCGCTGGCAGCTCTGAGGAGTATATATATACACCATCTTACAGCTAGACCTCTT 192
|||||

QY 61 TTGTAGAAAAGCAATGAGTGAAGTGCCTAAGT 96
|||||
Db 193 TTGTAGAAAAGCAATGAGTGAAGTGCCTAAGT 228
|||||

Search completed: May 2, 2003, 12:53:45
Job time : 467.145 secs

GenCore version 5.1.5
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 2, 2003, 11:21:20 ; Search time 152,938 Seconds
(without alignments)
1413.592 Million cell updates/sec

Title: US-09-719-554-3_COPY_3065_3160

Perfect score: 96
Sequence: 1 gccgcctgcgactcctcgaagg.....tgagtgagtgacatgaagt 96

Scoring table: IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Searched: 2185239 seqs, 112599159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database: N.GeneSeq_101002.*

1: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT.*
2: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT.*
3: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT.*
4: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT.*
5: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT.*
6: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT.*
7: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT.*
8: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT.*
9: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT.*
10: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT.*
11: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT.*
12: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT.*
13: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT.*
14: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT.*
15: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT.*
16: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT.*
17: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT.*
18: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT.*
19: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT.*
20: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.*
21: /SID52/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.*
22: /SID52/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.*
23: /SID52/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.*
24: /SID52/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	96	100.0	1321	20 AAX25655
2	96	100.0	1321	21 AAX59205
3	96	100.0	1326	21 ABN97928
4	96	100.0	1715	23 AAS73701
5	96	100.0	2372	20 AAX25664
6	96	100.0	2372	21 AAX59214
7	96	100.0	4349	23 AAS76475
8	96	100.0	7466	23 AAS68626
9	96	100.0	10499	21 ABN97929

10	96	100.0	56093	24 ABL61744
11	95.6	99.6	7582	20 AAX25655
12	95.6	99.6	7582	21 AAX59215
13	94.4	98.3	2938	20 AAX25656
14	94.4	98.3	2938	21 AAX59206
15	76.6	79.8	6394	23 AAS84210
16	70.4	73.3	1774	21 ABN97913
17	67.4	70.2	764	19 AAV43234
18	67.4	70.2	764	20 AAX29707
19	67.4	70.2	800	19 AAV43235
20	67.4	70.2	800	20 AAX29708
21	64.6	67.3	1003	21 AAX63827
22	64.6	67.3	1056	21 AAX59190
23	64.6	67.3	1186	20 AAX29726
24	64.6	67.3	1247	20 AAX29725
25	64.6	67.3	1489	21 ABN97932
26	64.6	67.3	1511	19 AAV43228
27	64.6	67.3	1511	20 AAX29706
28	64.6	67.3	1511	21 AAX59188
29	64.6	67.3	2009	21 AAX59189
30	64.6	67.3	2052	21 AAX63825
31	64.6	67.3	46340	21 ABN97978
32	62.4	65.0	1334	22 AAS63167
33	46.4	48.3	570	22 ABA61595
34	46.4	48.3	570	22 AAK09898
35	46.4	48.3	570	22 AAK35791
36	46.4	48.3	570	22 AAI41508
37	46.4	48.3	570	24 ABS10034
38	46.4	48.3	2633	22 ABA18794
39	46.4	48.3	32191	22 AAI94772
40	46.4	48.3	32191	22 AAI64117
41	46.4	48.3	44840	22 AAK71803
42	44.8	46.7	584	21 AAC81017
43	44.8	46.7	584	21 AAS99863
44	44.8	46.7	800	21 AAC81014
45	44.8	46.7	800	22 AAI67214

ALIGNMENTS

RESULT 1				
AXX25655	AAX25655 standard: CDNA to mRNA; 1321 BP.			
XX	AC	AAX25655;		
XX	XX	21-MAY-1999 (first entry)		
DT	XX			
XX	XX	Human endogenous retrovirus W clone cl.6A2.		
DE	XX			
XX	XX	Clone; human endogenous retrovirus; genome; autoimmune disease;		
KW	XX	multiple sclerosis; rheumatoid polyarthritits; insulin-dependent diabetes;		
KW	XX	disseminated lupus erythematosus; pregnancy; chromosomal marker; ss.		
XX	XX			
OS	XX	Human endogenous retrovirus.		
XX	XX			
PN	XX	WO9902696-A1.		
PD	XX	21-JAN-1999.		
XX	XX			
PF	XX	06-JUL-1998; 98WO-FR01442.		
PR	XX	07-JUL-1997; 97FR-0008815.		
XX	XX	(INNR) BIO MERIEUX.		
PI	XX	Beseme F, Blond JL, Bouton O, Mallet F, Mandrand B;		
XX	XX	WPI; 1999-120897/10.		
DR	XX			
XX	XX	New nucleic acid sequences from human endogenous retrovirus-W -		
PT	XX	expressed exclusively in placenta and useful in diagnosis and		

Colon adenocarcinoma
Complete human end
Human endogenous r
Human endogenous r
Gag and partial po
DNA encoding novel
Human retroviral s
Multiple sclerosis
Clone LB13 from MS
Multiple sclerosis
Clone LA15 from MS
Nucleotide sequenc
Endogenous retrov
Insert from pMT21C
Insert from pMT28C
Human retroviral s
Multiple sclerosis
Clone CL2 from MSR
Endogenous retrov
Nucleotide sequenc
Human retroviral s
Human purified sec
Human foetal liver
Human brain expres
Human bone marrow
Probe #10194 used
Human genome-deriv
Human nervous syst
Human excretory re
Human bladder rela
Human immune/haema
Human breast tumou
Breast tumour-spec
Human breast tumou
B311D CDNA sequenc

PT therapy of autoimmune disease, and abnormal or failed pregnancy
XX
PS Claim 1; Page 48-49; 106pp; French.
XX
CC This sequence represents clone cl.6n2 of the human endogenous retrovirus
CC (HERV) W genome. The nucleic acids, their fragments or peptides encoded
CC by them are markers of autoimmune disease (e.g. multiple sclerosis,
CC rheumatoid polyarthritis, disseminated lupus erythematosus, insulin-
CC dependent diabetes and related pathologies) and of abnormal or
CC unsuccessful pregnancy and can be used as chromosomal markers for
CC susceptibility to these conditions, or proximity markers of genes
CC associated with this susceptibility.
XX
SQ Sequence 1321 BP; 365 A; 352 C; 277 G; 327 T; 0 other;
Query Match 100.0%; Score 96; DB 20; Length 1321;
Best Local Similarity 100.0%; Pred. No. 6.4e-24;
Matches 96; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GCCGCTGGCACTCTCTGAGGAGTATTAATTATACACCATCTTACAGCTAGACTCTT 60
Db 717 GCCGCTGGCACTCTCTGAGGAGTATTAATTATACACCATCTTACAGCTAGACTCTT 776
QY 61 TTGTAGAAAAGCAAAATGAGTGAAGTCCATAAGT 96
Db 777 TTGTAGAAAAGCAAAATGAGTGAAGTCCATAAGT 812
RESULT 2
ID AAA59205 standard; DNA: 1321 BP.
XX
AC AAA59205;
XX
DT 07-NOV-2000 (first entry)
XX
DE 5' non coding and partial gag fragment of HERV-W from human genome.
XX
KW Autoimmune disease; retrovirus; human endogenous retrovirus W; HERV-W;
XX 9ag gene; pregnancy; multiple sclerosis; T cell proliferation; ss.
XX
OS Homo sapiens.
XX
PN WO200043521-A2.
XX
PD 27-JUL-2000.
XX
PE 21-JAN-2000; 2000WO-FR00144.
XX
PR 21-JAN-1999; 99FR-0000888.
XX
PA (INMR) BIO MERIEUX.
XX
PI Paranhos-Baccala G, Mallet F, Voisset C;
XX WPI; 2000-499229/44.
DR
PT New nucleic acid from human endogenous retrovirus, useful e.g. for
PT diagnosis of autoimmune disease and complications of pregnancy,
PT contains at least part of the gag gene -
XX
PS Disclosure; Page 42-43; 53pp; French.
XX
CC The present sequence represents an endogenous retrovirus, useful e.g. for
CC diagnosis of autoimmune disease, and is
CC integrated into the human genome. The fragment is originally derived
CC from a novel retrovirus, human endogenous retrovirus W (HERV-W). The
CC HERV-W retrovirus is associated with autoimmune disease, failure of
CC pregnancy or disorders of pregnancy. The nucleic acid fragment, or
CC proteins derived from it, are useful for diagnosis of autoimmune
CC disease (specifically multiple sclerosis) and for monitoring pregnancy.
CC The nucleic acid fragments may also be used for in situ labelling of
CC isolated chromosomes, while the transcription product can be used to

CC study or monitor T cell proliferation in vitro.
XX
SQ Sequence 1321 BP; 365 A; 352 C; 277 G; 327 T; 0 other;
XX
CC Query Match 100.0%; Score 96; DB 21; Length 1321;
CC Best Local Similarity 100.0%; Pred. No. 6.4e-24;
CC Matches 96; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GCCGCTGGCACTCTCTGAGGAGTATTAATTATACACCATCTTACAGCTAGACTCTT 60
Db 717 GCCGCTGGCACTCTCTGAGGAGTATTAATTATACACCATCTTACAGCTAGACTCTT 776
QY 61 TTGTAGAAAAGCAAAATGAGTGAAGTCCATAAGT 96
Db 777 TTGTAGAAAAGCAAAATGAGTGAAGTCCATAAGT 812
RESULT 3
ID ABR97928 standard; DNA: 1326 BP.
XX
AC ABR97928;
XX
DT 01-AUG-2002 (first entry)
XX
DE Human retroviral sequence gag.
XX
KW Autoimmune disease; HERV-7q; chromosome 7q; immunotherapy;
XX multiple sclerosis; ds.
XX
OS Human retrovirus.
XX
PN WO967395-A1.
XX
PD 29-DEC-1999.
XX
PE 23-JUN-1999; 99WO-FR01513.
XX
PR 23-JUN-1998; 98FR-0007920.
XX
PA (INRM) INSERM INST NAT SANTE & RECH MEDICALE.
XX
PI Alliel PM, Perlin J, Rieger F;
XX
DR WPI; 2000-160587/14.
XX
PT New nucleic acid sequences of human endogenous retrovirus, HERV-7q,
PT used for diagnosis, treatment and prevention of autoimmune and
PT neurological diseases -
XX
PS Claim 2; Page 129; 225pp; French.
XX
CC The present invention relates to new nucleic acid sequences of human
CC endogenous retrovirus, HERV-7q, which is located on chromosome 7q.
CC Regulatory elements associated with HERV-7q may alter expression of other
CC genes (even remote genes) on the same chromosome, inducing immunological
CC and/or neurological changes (which may be pathological or protective/
CC curative). HERV-7q peptides can be used to improve efficiency of the
CC immune response, e.g. in immunotherapy. HERV-7q peptides and their coding
CC sequences can be used in immunogenic or vaccinating compositions, for
CC protection against autoimmune diseases, particularly multiple sclerosis.
CC The peptides may also be used (by sequence comparison) to detect/identify
CC endogenous retroviruses that are abnormally expressed in cancer,
CC neuropathologies or other autoimmune diseases. The present sequence was
CC used to illustrate the invention.
XX
SQ Sequence 1326 BP; 425 A; 306 C; 303 G; 292 T; 0 other;
Query Match 100.0%; Score 96; DB 21; Length 1326;
Best Local Similarity 100.0%; Pred. No. 6.4e-24;
Matches 96; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GCCGCTGGCACTCTCTGAGGAGTATTAATTATACACCATCTTACAGCTAGACTCTT 60

|||||
Db 1 GCCGCTGGCACTCCTGAGGGAAGTATTAATTAACACCATCTTACAGCTAGACTCTT 60
Oy 61 TTGTAGAAAAGGCAATGAGTGAAGTGCCTAAGT 96
Db 61 TTGTAGAAAAGGCAATGAGTGAAGTGCCTAAGT 96

RESULT 4
AAS73701
ID AAS73701 standard; cDNA: 1715 BP.
XX
AC AAS73701;
XX
DT 13-FEB-2002 (first entry)
XX

DE DNA encoding novel human diagnostic protein #9505.
XX
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX
OS Homo sapiens.
XX
PN WO200175067-A2.
XX
PD 11-OCT-2001.
XX
PE 30-MAR-2001; 2001WO-US08631.
XX
PR 31-MAR-2000; 2000US-0540217.
PR 23-AUG-2000; 2000US-0649167.
XX
XX (HYSE-) HXSEQ INC.
XX
PI Drmanac RT, Liu C, Tang YT;
XX
DR WPI; 2001-639362/73.
DR P-PSDB; ABG09514.
XX
XX

PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity -
XX
XX

PS Claim 1; SEQ ID NO 9505; 103pp; English.
XX
XX

CC The invention relates to isolated polynucleotide (I) and
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. AAS64197-AAS94564 represent novel human
CC diagnostic coding sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
XX

Sequence 1715 BP; 512 A; 435 C; 371 G; 397 T; 0 other;

Query Match 100.0%; Score 96; DB 23; Length 1715;
Best Local Similarity 100.0%; Pred. No. 6; 9e-24;
Matches 96; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 GCCGCTGGCACTCCTGAGGGAAGTATTAATTAACACCATCTTACAGCTAGACTCTT 60
Db 655 GCCGCTGGCACTCCTGAGGGAAGTATTAATTAACACCATCTTACAGCTAGACTCTT 714
Oy 61 TTGTAGAAAAGGCAATGAGTGAAGTGCCTAAGT 96
Db 715 TTGTAGAAAAGGCAATGAGTGAAGTGCCTAAGT 750

RESULT 5
AAX25664
ID AAX25664 standard; cDNA to mRNA; 2372 BP.
XX
AC AAX25664;
XX
DT 21-MAY-1999 (first entry)
XX
DE Human endogenous retrovirus W clone cl.44.4.
XX
KW Clone; human endogenous retrovirus; genome; autoimmune disease;
KW multiple sclerosis; rheumatoid polyarthritis; insulin-dependent diabetes;
KW disseminated lupus erythematosus; pregnancy; chromosomal marker; ss.
XX
OS Human endogenous retrovirus.
XX
PN WO9902696-A1.
XX
PD 21-JAN-1999.
XX
PR 06-JUL-1998; 98WO-FR01442.
PF 07-JUL-1997; 97FR-0008815.
XX
XX (INMR) BIO MERIEUX.
XX
PA Beseme F, Blond JL, Bouton O, Mallet F, Mandrand B;
XX
PI WPI; 1999-120897/10.
XX
DR WPI; 1999-120897/10.
XX
XX

PT New nucleic acid sequences from human endogenous retrovirus-W -
PT expressed exclusively in placenta and useful in diagnosis and
PT therapy of autoimmune disease, and abnormal or failed pregnancy
XX
XX

PS Claim 1; Page 68-70; 106pp; French.
XX
XX

CC This sequence represents clone cl.44.4 of the human endogenous retrovirus
CC (HERV) W genome. The nucleic acids, their fragments or peptides encoded
CC by them are markers of autoimmune disease (e.g. multiple sclerosis,
CC rheumatoid polyarthritis, disseminated lupus erythematosus, insulin-
CC dependent diabetes and related pathologies) and of abnormal or
CC unsuccessful pregnancy and can be used as chromosomal markers for
CC susceptibility to these conditions, or proximity markers of genes
CC associated with this susceptibility.
XX
XX

Sequence 2372 BP; 674 A; 593 C; 544 G; 557 T; 4 other;

Query Match 100.0%; Score 96; DB 20; Length 2372;
Best Local Similarity 100.0%; Pred. No. 7; 3e-24;
Matches 96; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 GCCGCTGGCACTCCTGAGGGAAGTATTAATTAACACCATCTTACAGCTAGACTCTT 60
Db 603 GCCGCTGGCACTCCTGAGGGAAGTATTAATTAACACCATCTTACAGCTAGACTCTT 662

Oy 61 TTGTAGAAAAGGCAATGAGTGAAGTGCCTAAGT 96
Db 663 TTGTAGAAAAGGCAATGAGTGAAGTGCCTAAGT 698

RESULT 6
AAA59214
ID AAA59214 standard; DNA; 2372 BP.

XX AC AAA59214;
XX XX
DT 07-NOV-2000 (first entry)
XX XX
DE R-U5 region and partial pol gene sequences of HERV-W from human genome.
XX XX
KM Autoimmune disease; retrovirus; human endogenous retrovirus W; HERV-W;
XX gag gene; pregnancy; multiple sclerosis; T cell proliferation; ss.
XX OS Homo sapiens.
XX PN WO200043521-A2.
XX PN 27-JUL-2000.
XX PD 21-JAN-2000; 2000MO-FR00144.
XX PF 21-JAN-1999; 99FR-0000888.
XX PR (INMR) BIO MERIEUX.
XX PA Paranhos-Baccala G, Mallet F, Voisset C;
XX PI WPI; 2000-499229/44.
XX DR New nucleic acid from human endogenous retrovirus, useful e.g. for
XX PT diagnosis of autoimmune disease and complications of pregnancy,
XX PT contains at least part of the gag gene
XX PS disclosure; Page 49; 53pp; French.
XX XX
CC The present sequence represents an endogenous retroviral nucleic acid
CC fragment, which is associated with an autoimmune disease, and is
CC integrated into the human genome. The fragment is originally derived
CC from a novel retrovirus, human endogenous retrovirus W (HERV-W). The
CC HERV-W retrovirus is associated with autoimmune disease, failure of
CC pregnancy or disorders of pregnancy. The nucleic acid fragment, or
CC proteins derived from it, are useful for diagnosis of autoimmune
CC disease (specifically multiple sclerosis) and for monitoring pregnancy.
CC The nucleic acid fragments may also be used for in situ labelling of
CC isolated chromosomes, while the transcription product can be used to
CC study or monitor T cell proliferation in vitro.
XX XX
SQ Sequence 2372 BP; 674 A; 593 C; 544 G; 557 T; 4 other;
XX XX
Query Match 100.0%; Score 96; DB 21; Length 2372;
Best Local Similarity 100.0%; Pred. No. 7.5e-24;
Matches 96; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GCCGCGCTGGCAGCTCCGAGGAGATTAATTAATTAACACCATCTTACAGCTAGACTCTT 60
DB 603 GCCGCGCTGGCAGCTCCGAGGAGATTAATTAATTAACACCATCTTACAGCTAGACTCTT 662
QY 61 TTGTAGAAAAGCAATGAGTGAAGTGCATAGT 96
DB 663 TTGTAGAAAAGCAATGAGTGAAGTGCATAGT 698
RESULT 7
AAS76475
ID AAS76475 standard; cDNA; 4349 BP.
XX XX
AC AAS76475;
XX XX
DT 13-FEB-2002 (first entry)
XX XX
DE DNA encoding novel human diagnostic protein #12279.
XX XX
KM Human; chromosome mapping; gene mapping; gene therapy; forensic;
XX food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX OS Homo sapiens.

XX XX
PN WO200175067-A2.
XX XX
PD 11-OCT-2001.
XX XX
PF 30-MAR-2001; 2001MO-US08631.
XX XX
PR 31-MAR-2000; 2000US-0540217.
XX PR 23-AUG-2000; 2000US-0649167.
XX XX
PA (HYSE-) HYSEQ INC.
XX XX
PI Drmanac RT, Liu C, Tang YF;
XX XX
DR WPI; 2001-639362/73.
XX DR P-PSDB; ABG12288.
XX XX
PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity
XX PS Claim 1; SEQ ID No 12279; 103pp; English.
XX XX
CC The invention relates to isolated polynucleotide (I) and
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. AAS64197-AAS94564 represent novel human
CC diagnostic coding sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pat_sequences.
XX XX
SQ Sequence 4349 BP; 1319 A; 1061 C; 941 G; 1026 T; 2 other;
XX XX
Query Match 100.0%; Score 96; DB 23; Length 4349;
Best Local Similarity 100.0%; Pred. No. 8.8e-24;
Matches 96; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GCCGCGCTGGCAGCTCCGAGGAGATTAATTAATTAACACCATCTTACAGCTAGACTCTT 60
DB 608 GCCGCGCTGGCAGCTCCGAGGAGATTAATTAATTAACACCATCTTACAGCTAGACTCTT 667
QY 61 TTGTAGAAAAGCAATGAGTGAAGTGCATAGT 96
DB 668 TTGTAGAAAAGCAATGAGTGAAGTGCATAGT 703
RESULT 8
AAS68626
ID AAS68626 standard; cDNA; 7466 BP.
XX XX
AC AAS68626;
XX XX
DT 13-FEB-2002 (first entry)
XX XX
DE DNA encoding novel human diagnostic protein #4430.
XX XX
KM Human; chromosome mapping; gene mapping; gene therapy; forensic;
XX food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX XX

XX Homo sapiens.
OS
XX WO200175067-A2.
PN
XX 11-OCT-2001.
PD
XX 30-MAR-2001; 2001WO-US08631.
PE
XX 31-MAR-2000; 2000US-0540217.
PR
XX 23-AUG-2000; 2000US-0649167.
PA
XX (HYSE-) HYSEQ INC.
XX
XX Drmanac RT, Liu C, Tang YT;
PI
XX WPI: 2001-639362/73.
DR
XX P-PSDB: ABG04439.
XX
XX New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity
XX
XX Claim 1; SEQ ID No 4430; 103pp; English.
PS
XX The invention relates to isolated polynucleotide (I) and
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. AAS64197-AAS94564 represent novel human
CC diagnostic coding sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIFO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 7466 BP; 2217 A; 1880 C; 1614 G; 1754 T; 1 other;
Query Match 100.0%; Score 96; DB 23; Length 7466;
Best Local Similarity 100.0%; Pred. No. 1e-23;
Matches 96; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GCCGCTGGACCTCTGAGGAGATATAATTATACACATCTTACAGCTAGACCTCT 60
DB 4651 GCCGCTGGACCTCTGAGGAGATATAATTATACACATCTTACAGCTAGACCTCT 4710
QY 61 TTGTAGAAAAGGCAATGAGTGAAGTGCCTAAGT 96
DB 4711 TTGTAGAAAAGGCAATGAGTGAAGTGCCTAAGT 4746

RESULT 9
ABN97929
ID ABN97929 standard; DNA; 10499 BP.
XX
XX AC ABN97929;
XX
XX DT 01-AUG-2002 (first entry)
XX
XX DE Human retroviral sequence HERV-7q.
XX

KW Autoimmune disease; HERV-7q; chromosome 7q; immunotherapy;
XX multiple sclerosis; ds.
XX
XX Human retrovirus.
OS
XX WO9967395-A1.
PN
XX 29-DEC-1999.
PD
XX 23-JUN-1999; 99WO-FR01513.
PE
XX 23-JUN-1998; 98FR-0007920.
PR
XX (INRM) INSERM INST NAT SANTE & RECH MEDICALE.
PA
XX Alliel PM, Perin J, Rieger F;
PI
XX WPI: 2000-160587/14.
DR
XX
XX New nucleic acid sequences of human endogenous retrovirus, HERV-7q,
PT used for diagnosis, treatment and prevention of autoimmune and
PT neurological diseases
XX
XX Claim 3; Fig 1; 225pp; French.
PS
XX The present invention relates to new nucleic acid sequences of human
CC endogenous retrovirus, HERV-7q, which is located on chromosome 7q.
CC Regulatory elements associated with HERV-7q may alter expression of other
CC genes (even remote genes) on the same chromosome, inducing immunological
CC and/or neurological changes (which may be pathological or protective/
CC curative). HERV-7q peptides can be used to improve efficiency of the
CC immune response, e.g. in immunotherapy. HERV-7q peptides and their coding
CC sequences can be used in immunogenic or vaccinating compositions, for
CC protection against autoimmune diseases, particularly multiple sclerosis.
CC The peptides may also be used (by sequence comparison) to detect/identify
CC endogenous retroviruses that are abnormally expressed in cancer,
CC neuropathologies or other autoimmune diseases. The present sequence was
CC used to illustrate the invention.
XX
SQ Sequence 10499 BP; 3048 A; 2676 C; 2280 G; 2495 T; 0 other;
Query Match 100.0%; Score 96; DB 21; Length 10499;
Best Local Similarity 100.0%; Pred. No. 1.1e-23;
Matches 96; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GCCGCTGGACCTCTGAGGAGATATAATTATACACATCTTACAGCTAGACCTCT 60
DB 3065 GCCGCTGGACCTCTGAGGAGATATAATTATACACATCTTACAGCTAGACCTCT 3124
QY 61 TTGTAGAAAAGGCAATGAGTGAAGTGCCTAAGT 96
DB 3125 TTGTAGAAAAGGCAATGAGTGAAGTGCCTAAGT 3160

RESULT 10
ABL61744
ID ABL61744 standard; DNA; 56093 BP.
XX
XX ABL61744;
XX
XX DT 15-MAY-2002 (first entry)
XX
XX DE Colon adenocarcinoma related gene sequence SEQ ID NO:81.
XX
XX Human; cancer; colon; breast; ovary; oesophagus; kidney; thyroid;
KW stomach; lung; prostate; pancreas; carcinoma; antitumour; cancerous;
KW cytostatic; gene therapy; antineoplastic; Wilms tumour; adenocarcinoma;
KW gene; ds.
XX
XX OS Homo sapiens.
XX
XX PN WO200194629-A2.
XX

13-DEC-2001.
30-MAY-2001; 2001WO-US10838.
05-JUN-2000; 2000US-209473P.
05-JUN-2000; 2000US-209531P.
18-SEP-2000; 2000US-231133P.
18-SEP-2000; 2000US-233617P.
20-SEP-2000; 2000US-234009P.
20-SEP-2000; 2000US-234034P.
20-SEP-2000; 2000US-234052P.
22-SEP-2000; 2000US-234509P.
22-SEP-2000; 2000US-234567P.
25-SEP-2000; 2000US-234923P.
25-SEP-2000; 2000US-234924P.
25-SEP-2000; 2000US-235077P.
25-SEP-2000; 2000US-235082P.
25-SEP-2000; 2000US-235134P.
25-SEP-2000; 2000US-235280P.
26-SEP-2000; 2000US-235637P.
26-SEP-2000; 2000US-235638P.
27-SEP-2000; 2000US-235711P.
27-SEP-2000; 2000US-235720P.
27-SEP-2000; 2000US-235840P.
27-SEP-2000; 2000US-235863P.
28-SEP-2000; 2000US-236028P.
28-SEP-2000; 2000US-236032P.
28-SEP-2000; 2000US-236033P.
28-SEP-2000; 2000US-236034P.
28-SEP-2000; 2000US-236109P.
28-SEP-2000; 2000US-236111P.
29-SEP-2000; 2000US-236842P.
29-SEP-2000; 2000US-236891P.
29-SEP-2000; 2000US-236891P.
02-OCT-2000; 2000US-237172P.
02-OCT-2000; 2000US-237173P.
02-OCT-2000; 2000US-237278P.
02-OCT-2000; 2000US-237294P.
02-OCT-2000; 2000US-237295P.
02-OCT-2000; 2000US-237316P.
03-OCT-2000; 2000US-237425P.
03-OCT-2000; 2000US-237598P.
03-OCT-2000; 2000US-237604P.
03-OCT-2000; 2000US-237606P.
03-OCT-2000; 2000US-237608P.
01-NOV-2000; 2000US-244867P.
01-NOV-2000; 2000US-245084P.
(AVAL-) AVALON PHARM.
Young PE, Augustus M, Carter KC, Ebner R, Endress G, Horrigan S;
Soppet DR, Weaver Z;
WPI; 2002-188264/24.
Claim 1; SEQ ID 81; 44pp; English.
The present invention describes a method (M1) for screening for an anti-neoplastic agent. The method involves exposing cells to a chemical agent to be tested for anti-neoplastic activity, determining a change in expression of at least one gene (I) of a signature gene set, where (I) comprises a sequence (S) selected from 8447 sequences (given in ABL61666 to ABL70110), or is at least 95% identical to (S), where a change in expression is indicative of anti-neoplastic activity. (I) has cytoskeletal activity and can be used in gene therapy. M1 can be used for screening an anti-neoplastic agent, and can be used for producing a product which is the data collected with respect to the anti-neoplastic agent as a result of M1, and the data is sufficient to convey the chemical structure and/or properties of the agent. M1 can be used in the treatment of cancer such as colon, breast, stomach, lung, thyroid, CC

Query Match	100.0%	Score 96:	DB 24:	Length 56093;
Best Local Similarity	100.0%	Pred. No. 1.7e-23;		
Matches 96;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0
1	GGCGCTGGCATTCCGAGAGGAGATATATATACACCACTTACCTGACCTCTT	60		
Db	31065	GGCGCTGGCATTCCGAGAGGAGATATATATATACACCACTTACCTGACCTCTT	31124	
61	TTGTAGAAAAGCAATGGAGTGAAGTGCACCAATAGT	96		
Db	31125	TTGTAGAAAAGCAATGGAGTGAAGTGCACCAATAGT	31160	
RESULT 11				
AA25665				
ID	AA25665	standard; cDNA to mRNA; 7582 BP.		
AC	AA25665;			
XX				
DT	21-MAY-1999	(first entry)		
XX				
DE	Complete human endogenous retrovirus W genome.			
XX				
KW	Clone; human endogenous retrovirus; genome; autoimmune disease;			
RW	multiple sclerosis; rheumatoid polyarthritis; insulin-dependent diabetes;			
KW	disseminated lupus erythematosus; pregnancy; chromosomal marker; ss.			
XX				
OS	Human endogenous retrovirus.			
XX				
PN	W09902696-A1.			
XX				
PD	21-JAN-1999.			
XX				
PF	06-JUL-1998; 98WO-FR01442.			
XX				
PR	07-JUL-1997; 97FR-0008815.			
XX				
PA	(INMR) BIO MERIEUX.			
XX				
PI	Beseme F, Blond JL, Bouton O, Mallet F, Mandrand B;			
XX				
DR	WPI; 1999-120897/10.			
XX				
PT	New nucleic acid sequences from human endogenous retrovirus-W -			
PT	expressed exclusively in placenta and useful in diagnosis and			
PT	therapy of autoimmune disease, and abnormal or failed pregnancy			
XX				
PS	Claim 1; Page 71-74; 106pp; French.			
XX				
CC	This sequence represents the complete sequence of the human endogenous			
CC	retrovirus (HERV) W genome. The nucleic acids, their fragments or			
CC	peptides encoded by them are markers of autoimmune disease (e.g. multiple			
CC	sclerosis, rheumatoid polyarthritis, disseminated lupus erythematosus,			
CC	insulin-dependent diabetes and related pathologies) and of abnormal or			
CC	unsuccessful pregnancy and can be used as chromosomal markers for			
CC	susceptibility to these conditions, or proximity markers of genes			
CC	associated with this susceptibility.			
XX				
SO	Sequence 7582 BP; 2156 A; 1877 C; 1537 G; 1796 T; 2 U; 214 other:			
Query Match	99.6%;	Score 95.6;	DB 20;	Length 7582;
Best Local Similarity	99.0%;	Pred. No. 1.4e-23;		
Matches 95;	Conservative 1;	Mismatches 0;	Indels 0;	Gaps 0;
1	GGCGCTGGCATTCCGAGAGGAGATATATATACACCACTTACCTGACCTCTT	60		

```
DB 718 GCCGCTGGCAGCTCCTGAGGAGATATATAACACCATTTACAGCTAGACTCTT 777
OY 61 TTGTAGAAAAGCAATGAGTGAAGTGCCATAAGT 96
   |||||||
DB 778 TTGTAGAAAAGCAATGAGTGAAGTGCCATAAGT 813

RESULT 12
AAA59215
ID AAA59215 standard; DNA; 7582 BP.
AC AAA59215;
XX
XX 07-NOV-2000 (first entry)
DE Human endogenous retrovirus W (HERV-W) sequence.
XX
XX Autoimmune disease; retrovirus; human endogenous retrovirus W; HERV-W;
XX gag gene; pregnancy; multiple sclerosis; T cell proliferation; ss.
XX
XX Human endogenous retrovirus.
OS
FH Key Location/Qualifiers
FT LTR 1..120
   /tag= a
   /note= "R of 5' LTR"
FT LTR 121..575
   /tag= b
   /note= "U5 of 5' LTR"
FT primer_bind 579..596
   /tag= c
   /note= "581..7194"
FT CDS 581..7194
   /tag= d
   /note= "ORF1 env538"
FT CDS 7039..7194
   /tag= e
   /note= "ORF2 52 AA"
FT CDS 7112..7255
   /tag= f
   /note= "ORF3 48 AA"
FT misc_feature 7244..7254
   /tag= g
   /note= "polyurine tract"
FT LTR 7256..7582
   /tag= h
   /note= "U3-R of 3' LTR"
FT polyA_signal 7563..7569
   /tag= i
   /note= "1"
XX PN WO200043521-A2.
XX PD 27-JUL-2000.
XX PF 21-JAN-2000; 2000WO-FR00144.
XX PR 21-JAN-1999; 99FR-0000888.
XX PA (INMR ) BIO MERIEUX.
XX PI Paranhos-Baccala G, Mallet F, Voliset C;
XX WPI; 2000-499229/44.
XX DR WPI; 2000-499229/44.
XX PT New nucleic acid from human endogenous retrovirus, useful e.g. for
XX diagnosis of autoimmune disease and complications of pregnancy,
XX contains at least part of the gag gene
XX
XX PS Disclosure; Page 49-52; 53pp; French.
XX
XX The present sequence represents an endogenous retrovirus, which is
XX associated with an autoimmune disease, and is integrated into the human
XX genome. The retrovirus is human endogenous retrovirus W (HERV-W). The
XX HERV-W retrovirus is associated with autoimmune disease, failure of
```

```
CC pregnancy or disorders of pregnancy. HERV-W nucleic acid fragments, or
CC proteins derived from it, are useful for diagnosis of autoimmune
CC disease (specifically multiple sclerosis) and for monitoring pregnancy.
CC The nucleic acid fragments may also be used for in situ labelling of
CC isolated chromosomes, while the transcription product can be used to
CC study or monitor T cell proliferation in vitro.
XX
SQ Sequence 7582 BP; 2156 A; 1876 C; 1538 G; 1796 T; 216 other;
Query Match 99.6%; Score 95.6; DB 21; Length 7582;
Best Local Similarity 99.0%; Pred. No. 1.4e-23;
Matches 95; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
OY 1 GCCGCTGGCAGCTCCTGAGGAGATATATAACACCATTTACAGCTAGACTCTT 60
   |||||||
DB 718 GCCGCTGGCAGCTCCTGAGGAGATATATAACACCATTTACAGCTAGACTCTT 777
OY 61 TTGTAGAAAAGCAATGAGTGAAGTGCCATAAGT 96
   |||||||
DB 778 TTGTAGAAAAGCAATGAGTGAAGTGCCATAAGT 813
```

RESULT 13

AAK25656

ID AAK25656 standard; CDNA to mRNA; 2938 BP.

XX AAK25656;

XX 21-MAY-1999 (first entry)

XX Human endogenous retrovirus W clone cl.6A1.

XX Clone; human endogenous retrovirus; genome; autoimmune disease;

XX multiple sclerosis; rheumatoid polyarthritis; insulin-dependent diabetes;

XX disseminated lupus erythematosus; pregnancy; chromosomal marker; ss.

XX Human endogenous retrovirus.

XX WO9902696-A1.

XX 21-JAN-1999.

XX 06-JUL-1998; 98WO-FR01442.

XX 07-JUL-1997; 97FR-0008815.

XX (INMR) BIO MERIEUX.

XX Beseme F, Blond JL, Bouton O, Mallet F, Mandrand B;

XX WPI; 1999-120897/10.

XX New nucleic acid sequences from human endogenous retrovirus-W -

XX expressed exclusively in placenta and useful in diagnosis and

XX therapy of autoimmune disease, and abnormal or failed pregnancy

XX Claim 1; Page 49-52; 106pp; French.

XX This sequence represents clone cl.6A1 of the human endogenous retrovirus

XX (HERV) W genome. The nucleic acids, their fragments or peptides encoded

XX by them are markers of autoimmune disease (e.g. multiple sclerosis,

XX rheumatoid polyarthritis, disseminated lupus erythematosus, insulin-

XX dependent diabetes and related pathologies) and of abnormal or

XX unsuccessful pregnancy and can be used as chromosomal markers for

XX susceptibility to these conditions, or proximity markers of genes

XX associated with this susceptibility.

SQ Sequence 2938 BP; 878 A; 720 C; 646 G; 692 T; 2 other;

Query Match 98.3%; Score 94.4; DB 20; Length 2938;

Best Local Similarity 99.0%; Pred. No. 2.9e-23;

Matches 95; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 GCCGCTGGACCTCTGAGGAGATATAATTAACACCATCTTAGAGCTAGACCTCTT 60
DB 133 GCCGCTGGACCTCTGAGGAGATATAATTAACACCATCTTAGAGCTAGACCTCTT 192
OY 61 TTGTAGAAAAGGCAATGAGTGAAGTCCATTAAGT 96
DB 193 TTGTAGAAAAGGCAATGAGTGAAGTCCATTAAGT 228

RESULT 14

AAA59206
ID AAA59206 standard; DNA; 2938 BP.

AAA59206;

07-NOV-2000 (first entry)

XX Gag and partial pol gene fragment of HENV-W from human genome.

XX Autoimmune disease; retrovirus; human endogenous retrovirus W; HENV-W;

XX gag gene; pregnancy; multiple sclerosis; T cell proliferation; ss.

OS Homo sapiens.

XX MO200043521-A2.

XX 27-JUL-2000.

XX 21-JAN-2000; 2000WO-FR00144.

XX 21-JAN-1999; 99FR-0000888.

XX (1NMR) BIO MERIEUX.

XX Paranhos-Baccala G, Mallet F, Voliset C;

XX WPI: 2000-499229/44.

XX New nucleic acid from human endogenous retrovirus; useful e.g. for

XX diagnosis of autoimmune disease and complications of pregnancy,

XX contains at least part of the gag gene -

XX Disclosure; Page 43; 53pp; French.

XX The present sequence represents an endogenous retroviral nucleic acid

XX fragment, which is associated with an autoimmune disease, and is

XX integrated into the human genome. The fragment is originally derived

XX from a novel retrovirus, human endogenous retrovirus W (HENV-W). The

XX HENV-W retrovirus is associated with autoimmune disease, failure of

XX pregnancy or disorders of pregnancy. The nucleic acid fragment, or

XX proteins derived from it, are useful for diagnosis of autoimmune

XX disease (specifically multiple sclerosis) and for monitoring pregnancy.

XX The nucleic acid fragments may also be used for in situ labelling of

XX isolated chromosomes, while the transcription product can be used to

XX study or monitor T cell proliferation in vitro.

XX Sequence 2938 BP; 878 A; 720 C; 646 G; 692 T; 2 other;

Query Match

Best Local Similarity 98.3%; Score 94.4; DB 21; Length 2938;

Matches 95; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

ID AAS84210 standard; CDNA; 6394 BP.
XX AAS84210;
AC AAS84210;
XX 13-FEB-2002 (first entry)
DT 13-FEB-2002 (first entry)
DE DNA encoding novel human diagnostic protein #20014.
XX Human; chromosome mapping; gene mapping; gene therapy; forensic;
XX food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX Homo sapiens.
XX MO200175067-A2.
XX 11-OCT-2001.
XX 30-MAR-2001; 2001WO-US08631.
XX 31-MAR-2000; 2000US-0540217.
XX 23-AUG-2000; 2000US-0649167.
XX (HXSE-) HXSEQ INC.
XX Drmanac RT, Liu C, Tang YT;
XX WPI: 2001-639362/73.
XX P-PSDB: ABG20023.

XX New isolated polynucleotide and encoded polypeptides, useful in

XX diagnostics, forensics, gene mapping, identification of mutations

XX responsible for genetic disorders or other traits and to assess

XX biodiversity -

XX Claim 1; SEQ ID NO 20014; 103pp; English.

XX The invention relates to isolated polynucleotide (I) and

XX polypeptide (II) sequences. (I) is useful as hybridisation probes,

XX polymerase chain reaction (PCR) primers, oligomers, and for chromosome

XX and gene mapping, and in recombinant production of (II). The

XX polynucleotides are also used in diagnostics as expressed sequence tags

XX for identifying expressed genes. (I) is useful in gene therapy techniques

XX to restore normal activity of (II) or to treat disease states involving

XX (II). (II) is useful for generating antibodies against it, detecting or

XX quantitating a polypeptide in tissue, as molecular weight markers and as

XX a food supplement. (II) and its binding partners are useful in medical

XX imaging of sites expressing (II). (I) and (II) are useful for treating

XX disorders involving aberrant protein expression or biological activity.

XX The polypeptide and polynucleotide sequences have applications in

XX diagnostics, forensics, gene mapping, identification of mutations

XX responsible for genetic disorders or other traits to assess biodiversity

XX and to produce other types of data and products dependent on DNA and

XX amino acid sequences. AAS64197-AAS94564 represent novel human

XX diagnostic coding sequences of the invention.

XX Note: The sequence data for this patent did not appear in the printed

XX specification, but was obtained in electronic format directly from WIPO

XX at ftp.wipo.int/pub/published_pcl_sequences.

XX Sequence 6394 BP; 1840 A; 1597 C; 1384 G; 1571 T; 2 other;

XX Query Match 79.8%; Score 76.6; DB 23; Length 6394;

XX Best Local Similarity 92.9%; Pred. No. 6.6e-17;

XX Matches 92; Conservative 0; Mismatches 4; Indels 3; Gaps 1;

OY 1 GCCGCTGGACCTCTGAGGAGATATAATTAACACCATCTTAGAGCTAGACCTCTT 60

DB 2697 GCCGCTGGACCTCTGAGGAGATATAATTAACACCATCTTAGAGCTAGACCTCTT 2756

OY 61 TTGTAGAAA--AAGCAATGAGTGAAGTCCATTAAGT 96

DB 2757 TTGTAGAAAAGGCAATGAGTGAAGTCCATTAAGT 2795

RESULT 15

AAS84210

Mon May 5 08:41:35 2003

us-09-719-554-3_copy_3065_3160.rng

Page 9

Search completed: May 2, 2003, 11:29:57
Job time : 172.938 secs

GenCore version 5.1.5
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 2, 2003, 11:21:20 ; Search time 1105.66 seconds

(Without alignments)
1406.195 Million cell updates/sec

Title: US-09-719-554-3_COPY_3065_3160

Perfect score: 96
Sequence: 1 gccgcctgcgcactcctgagg.....tggagtgagtgccataagt 96

Scoring table: IDENTITY_NUC
Gapop:10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database :

Listing first 45 summaries

EST: *
1: em_estba: *
2: em_esthum: *
3: em_estin: *
4: em_estnu: *
5: em_estov: *
6: em_estpl: *
7: em_estro: *
8: em_estc: *
9: gb_estl: *
10: gb_estl2: *
11: gb_estl3: *
12: gb_estl4: *
13: gb_estl5: *
14: gb_estl6: *
15: em_estfun: *
16: em_estom: *
17: gb_gss: *
18: em_gss_hum: *
19: em_gss_inv: *
20: em_gss_pln: *
21: em_gss_vit: *
22: em_gss_fun: *
23: em_gss_mam: *
24: em_gss_mus: *
25: em_gss_other: *
26: em_gss_pro: *
27: em_gss_fod: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	96	100.0	450	B0024410
2	68.8	71.7	913	B0424090
3	67.2	70.0	681	AG009944
4	67.2	70.0	685	AG009944
5	67.2	70.0	689	AG009943
6	67.2	70.0	692	AG009942

7	66.6	69.4	711	AG009923
8	64.6	67.3	401	AG014538
9	62.4	65.0	690	BM980860
10	61.4	64.0	559	AG083457
11	49.6	51.7	400	HS_5296_B
12	48	50.0	354	AG058957
13	48	50.0	373	AG010624
14	48	50.0	851	AG0748718
15	47.6	49.6	632	AG000365
16	46.4	48.3	429	AG0035279
17	46.4	48.3	559	AL708722
18	46	47.9	237	BA5372
19	44.8	46.7	545	AG0133205
20	44.8	46.7	711	AV730914
21	43.4	45.2	420	AG0435696
22	43.4	45.2	473	AG0248325
23	43.2	45.0	444	B99733
24	43.2	45.0	493	B91430
25	42.8	44.6	358	B03917
26	42.6	44.4	513	AG034931
27	42.6	44.4	645	AG18617
28	42.2	44.0	500	B53643
29	42.2	44.0	500	B53643
30	41.8	43.5	318	B53275
31	41.6	43.3	291	AA904177
32	41.6	43.3	314	AA983698
33	41.6	43.3	358	AT1219949
34	41.6	43.3	370	AM105235
35	41.6	43.3	378	BF960517
36	41.6	43.3	408	B98887
37	41.6	43.3	455	B60060
38	41.6	43.3	464	BM718032
39	41.6	43.3	465	BM680485
40	41.6	43.3	543	B71000
41	41.6	43.3	568	B55061
42	41.6	43.3	605	B55062
43	41.6	43.3	632	AG0021694
44	41	42.7	584	AG031335
45	40.6	42.3	477	AG024121

ALIGNMENTS

RESULT 1
B0024410/c
LOCUS
DEFINITION
B0024410
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

450 bp mRNA linear EST 27-MAR-2002
UI-1-BBIP-aug-e-06-0-UI.s1 NCI-CCAP P16 Homo sapiens CDNA clone
UI-1-BBIP-aug-e-06-0-UI 3', mRNA sequence.
B0024410
B0024410.1 GI:19759689
EST.
human.
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
NCI-CCAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgaabs-remail.nih.gov
Tissue Procurement: Dr. Steven Brown
CDNA Library preparation: Dr. M. Bento Soares, University of Iowa
CDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Clone distribution information can be obtained
from Dr. M. Bento Soares, bento-soares@uiowa.edu
The following repetitive elements were found in this CDNA
sequence: 14-226, >HERV17#LTR/Retroviral (matched complement)
214-256, >HERV17#LTR/Retroviral (matched complement) 257-450,
>LTR17#LTR/Retroviral (matched complement)

Seq primer: M13 FORWARD
POLYA-Yes.
Location/Qualifiers
1. 450
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="NCI-CCAP_P16"
/clone_1lb="NCI-CCAP_P16"
/tissue_type="Placenta"
/dev_stage="Full Term"
/lab_host="DH10B (Life Technologies)"
/note="Organ: Placenta; Vector: pT73-Pac (Pharmacia) with a modified polylinker; Site_1: EcoR I; Site_2: Not I; NCI-CCAP_P16 is a subtracted cDNA library constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was ligated to an EcoR I adaptor, digested with Not I, and cloned directionally into pT73-Pac vector. The oligonucleotide used to prime the synthesis of first-strand cDNA contains a library tag sequence that is located between the Not I site and the (dT)18 tail. The additional information, contact: Bento Soares, bento-soares@uiowa.edu
TAG_L1B=UI-1-BB1P
TAG_TISSUE=Placenta human full term
TAG_SEQ=AGGAA"
BASE COUNT 94 a 96 c 113 g 147 t
ORIGIN
Query Match 100.0%; Score 96; DB 14; Length 450;
Best Local Similarity 100.0%; Pred. No. 1.2e-20;
Matches 96; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GCCGCTGGACCTCTGAGGAGTATTAATTTACACCATCTTACAGTACCTCTT 60
DB 114 GCCGCTGGACCTCTGAGGAGTATTAATTTACACCATCTTACAGTACCTCTT 55
QY 61 TTGTAGAAAAGCAATGAGTGAAGTGCCTAAGT 96
DB 54 TTGTAGAAAAGCAATGAGTGAAGTGCCTAAGT 19
RESULT 2
BQ424090 913 bp mRNA linear EST 23-MAY-2002
LOCUS
DEFINITION AGNCOURT_7892347 NIH_MGC_72 Homo sapiens cDNA clone IMAGE:6158499
5', mRNA sequence.
ACCESSION BQ424090
VERSION BQ424090.1 GI:21119405
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 913)
NIH-MGC http://mgi.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabbs@mail.nih.gov
Tissue Procurement: ATCC/DC/DTP
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLM13506 row: b column: 04
High quality sequence stop: 619.
Location/Qualifiers
1. 913
FEATURES
source

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:6158499"
/clone_1lb="NIH_MGC_72"
/tissue_type="melanotic melanoma"
/lab_host="DH10B (Phage-resistant)"
/note="Organ: skin; Vector: pCMV-Sport6; Site_1: NotI; Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT. Average insert size 2 kb. Library constructed by Life Technologies."
BASE COUNT 238 a 267 c 196 g 212 t
ORIGIN
Query Match 71.7%; Score 68.8; DB 14; Length 913;
Best Local Similarity 94.3%; Pred. No. 9.2e-12;
Matches 83; Conservative 0; Mismatches 2; Indels 3; Gaps 1;
QY 12 CTCCTGAGGAGTATTAATTTATACACCATCTTACAGTACCTTTGTAGAA--- 68
DB 634 CTCCTGAGGAGTATTAATTTATACACCATCTTACAGTACCTTTGTAGAAAG 693
QY 69 AAGGCAATGAGTGAAGTGCCTAAGT 96
DB 694 AAGGCAATGAGTGAAGTGCCTAAGT 721
RESULT 3
AG009944/c 681 bp DNA linear GSS 14-APR-1999
LOCUS
DEFINITION Homo sapiens genomic DNA, 21q region, clone: 72018A32, genomic survey sequence.
ACCESSION AG009944 AG003755
VERSION AG009944.1 GI:3289930
KEYWORDS GSS.
SOURCE Homo sapiens DNA, clone: 72018A32.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 681)
Hattori, M., Ishii, K., Toyoda, A., Shiba, T. and Sakaki, Y.
Homo sapiens genomic DNA, chromosome 21q
Published Only in Database (1998)
2 (bases 1 to 681)
Hattori, M., Ishii, K., Toyoda, A., Shiba, T. and Sakaki, Y.
Direct Submission
Submitted (06-JUL-1998) Masahira Hattori, Kitasato University,
Department of Science, JST Sequencing Laboratory, Kitasato 1-15-1,
Sagamihara 228, Japan (E-mail: hattori@hgc.ims.u-tokyo.ac.jp,
Tel:0427-78-9732, Fax:0427-78-9561)
On Feb 5, 1999 this sequence version replaced g1:2754657.
AG003755: Submitted (07-Jan-1998).
Location/Qualifiers
1. 681
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="21"
/map="21q"
/clone="72018A32"
BASE COUNT 141 a 143 c 180 g 194 t 23 others
ORIGIN
Query Match 70.0%; Score 67.2; DB 17; Length 681;
Best Local Similarity 93.2%; Pred. No. 2.8e-11;
Matches 82; Conservative 0; Mismatches 3; Indels 3; Gaps 1;
QY 12 CTCCTGAGGAGTATTAATTTATACACCATCTTACAGTACCTTTGTAGAA---A 68
DB 442 CCCCTGAGGAGTATTAATTTATACACCATCTTACAGTACCTTTGTAGAAAGA 383
QY 69 AAGGCAATGAGTGAAGTGCCTAAGT 96
DB 382 AAGGCAATGAGTGAAGTGCCTAAGT 355
FEATURES
source

[illegible]

FEATURES	AG003754: Submitted (07-Jan-1998).					
SOURCE	Location/Qualifiers 1..688					
	/organism="Homo sapiens"					
	/db_xref="taxon:9606"					
	/chromosome="21"					
	/map="21q"					
	/clone="72018A32"					
BASE COUNT	145 a	145 c	188 g	198 t	13 others	
ORIGIN						
Query Match	70.0%;	Score 67.2;	DB 17;	Length 689;		
Best Local Similarity	93.2%;	Pred. No. 2.8e-11;				
Matches 82:	Conservative	0;	Mismatches 3;	Indels 3;	Gaps 1;	
OY	12	CTCCGAGGGAAGTATAATATTATACACCACCTTTACACTGAGCCTCTTTGTAGA--A	68			
Db	173	CCCCGAGGGAAGTATAATATTATACACCACCTTTACACTGAGCCTCTTTGTAGAAGA	114			
OY	69	AAGGCAATGGAGTGAGTAAGTCGCATTAAGT	96			
Db	113	AGGCGAAATGGAGTGAGTAAGTCGCATTAAGT	86			
RESULT 6						
LOCUS	AG009942/c	692 bp	DNA	linear	GSS 14-Apr-1999	
DEFINITION	Homo sapiens genomic DNA, 21q region, clone: 72018A32, genomic survey sequence.					
ACCESSION	AG009942 AG003753					
VERSION	AG009942.1 GI:3289928					
KEYWORDS	GSS.					
SOURCE	Homo sapiens					
ORGANISM	Homo sapiens DNA, clone:72018A32.					
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.					
AUTHORS	1 (bases 1 to 692)					
TITLE	Hattori,M., Ishii,K., Toyoda,A., Shiba,T. and Sakaki,Y.					
JOURNAL	Homo sapiens genomic DNA, chromosome 21q					
REFERENCE	Published Only In Database (1998)					
AUTHORS	2 (bases 1 to 692)					
TITLE	Hattori,M., Ishii,K., Toyoda,A., Shiba,T. and Sakaki,Y.					
JOURNAL	Direct Submission					
COMMENT	Submitted (06-JUL-1998) Masahira Hattori, Kitasato University, Department of Science, JST Sequencing Laboratory; Kitasato 1-15-1, Sagami-hara 228, Japan (E-mail:hattori@ngc.ims.u-tokyo.ac.jp, Tel:0427-78-9732, Fax:0427-78-9561)					
COMMENT	On Feb 5, 1999 this sequence version replaced gi:2754655.					
COMMENT	AG003753: Submitted (07-Jan-1998).					
FEATURES	Location/Qualifiers 1..692					
SOURCE	/organism="Homo sapiens"					
	/db_xref="taxon:9606"					
	/chromosome="21"					
	/map="21q"					
	/clone="72018A32"					
BASE COUNT	158 a	158 c	186 g	181 t	9 others	
ORIGIN						
Query Match	70.0%;	Score 67.2;	DB 17;	Length 692;		
Best Local Similarity	93.2%;	Pred. No. 2.8e-11;				
Matches 82:	Conservative	0;	Mismatches 3;	Indels 3;	Gaps 1;	

		Email: jwallace@u.washington.edu
		Clones are derived from the human BAC library RPCT-11. For BAC
		library availability, please contact Pieter de Jong
		(pietere@jorg.med.buffalo.edu). Clones may be purchased from
		BACPAC Resources (http://bacpac.med.buffalo.edu/ordering_bac.htm)
		or from Resear h Genetics (info@resgen.com). BAC end Web Server:
		http://www.hsc.washington.edu
		Plate: 701 row: B column: 6
		Seq primer: SP6
		Class: BAC ends
		High quality sequence stop: 401.
FEATURES		
Source		Location/Qualifiers
		1..401
		/organism="Homo sapiens"
		/db_xref="taxon:9606"
		/clone="plate=701 col=6 Row=B"
		/clone_id="RPCT-11 Human Male BAC Library"
		/sex="male"
		/note="Vector: pBACE3.6; Site_1: EcoRI; Site_2: EcoRI;
		Male blood DNA was isolated from one randomly chosen donor
		and partially digested with a combination of EcoRI and
		EcoRI Methylase. Size selected DNA was cloned into the
		pBACE3.6 vector at EcoRI sites"
BASE COUNT	95 a	77 c 87 g 140 t 2 others
ORIGIN		
Query Match	67.3%	Score 64.6; DB 17; Length 401;
Best Local Similarity	90.9%; Pred.No.1.6e-10;	
Matches 80; Conservative	0; Mismatches 5;	Indels 3; Gaps 1;
OY 12 CTCCTGAGGAAGTATTAATTATAACCACTTTACTAGACCCTTGTTGTAGA---A 68 Db 325 CCCTNGAGGAAGTATTAATTATAAACCACTTTACTAGACCCTTGTTGTAGAAGA 266 		
OY 69 AAGCGAATGGAGTAGTCGCATMACT 96 Db 265 AGGCCAAATGGAGTAGTCGCATMTGT 238 		
RESULT 9		
BM980860/c	690 bp - mRNA linear EST 21-MAR-2002	
LOCUS	BM980860	
DEFINITION	UI-CF-ENI-age-g-24-0-UT.51 UI-CF-ENI Homo sapiens CDNA clone	
ACCESSION	BM980860 GI:19602750	
VERSION	EST.	
KEYWORDS	human.	
SOURCE	Homo sapiens	
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo. 1 (bases 1 to 690) Ronald O.M.F., Lennon,G. and Soares,M.B. Normalization and subtraction: two approaches to facilitate gene discovery Genome Res. 6 (9), 791-806 (1996)	
REFERENCE	97044477	
AUTHORS	Contact: McCray, PB	
TITLE	McCrays Lab	
JOURNAL	University of Iowa	
MEDLINE	2024 University of Iowa Med Labs, Iowa City, IA 52242, USA	
COMMENT	Tel.: 319 356 4866 Fax: 319 356 7171 Email: paul.mccray@uiowa.edu Tissue Procurement: Dr. M. J. Welsh, University of Iowa CDNA Library Preparation: Dr. M. Bento Soares, University of Iowa DNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa DNA Sequencing by: Dr. M. Bento Soares, University of Iowa Clone Distribution: Researchers may obtain clones from Research Genetics (www.resgen.com).	
	The following repetitive elements were found in this CDNA sequence: 10-148, >HERV1/LTR/Retroviral (matched complement) 18-266, >HEBV9/LTR/Retroviral (matched complement) 176-296,	

>HERV17#LTR/Retroviral (matched compliment) 327-690,
>LTR17#LTR/Retroviral (matched compliment)
Seq primer: M13 FORWARD

POLYA=yes.

FEATURES

SOURCE

Location/Qualifiers
1. 690
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="UT-CF-EN1-ade-g-24-0-UT"
/clone_lib="UT-CF-EN1"
/tissue_type="Primary Lung Cystic Fibrosis Epithelial Cells"
/dev_stage="Adult"
/lab_host="DH10B (Life Technologies) (T1 phage resistant)"
/note="Organ: Lung; Vector: pT73-Pac (Pharmacia) with a modified polylinker; Site_1: EcoR I; Site_2: Not I; UI-CF-EN1 is a normalized cDNA library containing the following tissue(s): Primary Lung Cystic Fibrosis Epithelial Cells. The library was constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was ligated to an EcoR I adaptor, digested with Not I, and cloned directionally into pT73-Pac vector. The oligonucleotide used to prime the synthesis of the first-strand cDNA contains a library tag sequence that is located between the Not I site and the (dT)18 tail. The sequence tag for this library is CTGCTCAGGT.
TAG_L1B-UT-CF-EN1
TAG_TISSUE-Human Lung Epithelial Cell lines untreated LPS 6hr to LPS 24h
TAG_SEQ-CTGCTCAGGT"

BASE COUNT

151 a 155 c 191 g 193 t

ORIGIN

Query Match 65.0%; Score 62.4; DB 14; Length 690;
Best Local Similarity 89.8%; Pred. No. 1e-09;
Matches 79; Conservative 0; Mismatches 6; Indels 3; Gaps 1;

QY 12 CTCTGAGGGAATTAATTAATTAACACCACTTACAGCTTCTTTGTAGAA--- 68

DB 142 CTCTGAGGGAATTAATTAATTAACACCACTTACAGCTTCTTTGTAGAAAG 83

QY 69 AAGCAATGAGTGAAGTGCATATGT 96

DB 82 AAGCAATGAGTGAAGTGCATATGT 55

RESULT 10

LOCUS A0833457

DEFINITION HS_5596_B2_H12_T7A RPT-11 Human Male BAC Library Homo sapiens

ACCESSION A0833457

VERSION A0833457.1 GI:5799442

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

COMMENT

Sequence-tagged connectors: A sequence approach to mapping and scanning the human genome
Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)
Contact: Mahairas GG, Wallace JC, Hood L
High Throughput Sequencing Center
University of Washington
401 Queen Anne Avenue North, Seattle, WA 98109, USA
Tel: (206) 616-3618

Fax: (206) 616-3887

Email: jwallace@u.washington.edu

Clones are derived from the human BAC library RPT-11. For BAC library availability, please contact Pieter de Jong (pieterdejong.med.bufileo.edu). Clones may be purchased from BACPAC Resources (http://bacpac.med.bufileo.edu/ordering_bac.htm) or from Research Genetics (info@resgen.com). BAC end Web Server: <http://www.htsc.washington.edu>

Plate: 872 Row: P Column: 24

Seq primer: T7

Class: BAC ends

High quality sequence stop: 559.

Location/Qualifiers

1. 559

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="Plate=872 Col=24 Row=P"
/clone_lib="RPT-11 Human Male BAC Library"
/sex="male"

/note="Vector: pBAC3.6; Site_1: EcoRI; Site_2: EcoRI; Male blood DNA was isolated from one randomly chosen donor and partially digested with a combination of EcoRI and EcoRI Methylase. Size selected DNA was cloned into the pBAC3.6 vector at EcoRI sites"

161 a 146 c 117 g 122 t 13 others

BASE COUNT

ORIGIN

Query Match 64.0%; Score 61.4; DB 17; Length 559;
Best Local Similarity 88.6%; Pred. No. 2e-09;
Matches 78; Conservative 0; Mismatches 7; Indels 3; Gaps 1;

QY 12 CTCTGAGGGAATTAATTAATTAACACCACTTACAGCTTCTTTGTAG---AA 68

DB 328 CTCTGAGGGAATTAATTAATTAACACCACTTACAGCTTCTTTGTAGAGAG 387

QY 69 AAGCAATGAGTGAAGTGCATATGT 96

DB 388 AAGCAATGAGTGAAGTGCATATGT 415

RESULT 11

LOCUS T59288

DEFINITION Y502c10.r1 Striatagene Lung (#937210) Homo sapiens cDNA clone

ACCESSION T59288

VERSION T59288.1 GI:661125

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 400)
Hillier, L., Lennon, G., Becker, M., Bonaldo, M.F., Chiapelli, B., Chissoe, S., Dietrich, N., Dubuque, T., Favello, A., Gish, W., Hawkins, M., Hultman, M., Kucaba, T., Lacy, M., Le, M., Le, N., Mardis, E., Moore, B., Morris, M., Parsons, J., Prange, C., Rifkin, L., Rohlfing, T., Schellenberg, K., Soares, M.B., Tan, F., Thierly-Meg, J., Trevisan, E., Underwood, K., Wohlmann, P., Waterston, R., Wilson, R. and Marra, M.
Generation and analysis of 280,000 human expressed sequence tags
Genome Res. 6 (9), 807-828 (1996)

97044478

Contact: Willson RK

Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108

Tel: 314 286 1800

Fax: 314 286 1810

Email: est@watson.wustl.edu

High quality sequence stops: 309 Source: IMAGE Consortium, LINT This clone is available royalty-free through LINT; contact the IMAGE Consortium (info@image.lnl.gov) for further information.
Insert length: 908
Std Error: 0.00

Db	52	AAGGCAATGCAGTGAATACCTTATC	25
RESULT 14			
LOCUS	AQ748718/c		
DEFINITION	HS.554.O A2.B03.SP6.RPci-11 Human Male BAC library Homo sapiens genomic clone Plate=1116 Col=6 Row=C, DNA sequence.	851 bp	DNA linear GSS 19-JUL-1999
ACCESSION	AQ748718		
VERSION	AQ748718.1	GI:5535964	GSS.
KEYWORDS			
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo. 1 (bases 1 to 851)		
AUTHORS	Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T., Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and Hood,L..		
TITLE	Sequence-tagged connectors: A sequence approach to mapping and scanning the human genome		
JOURNAL	Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)		
MEDLINE	99380589		
COMMENT	Contact: Mahairas GG, Wallace JC, Hood L High Throughput Sequencing Center University of Washington 401 Queen Anne Avenue North, Seattle, WA 98109, USA Tel: (206) 616-3618 Fax: (206) 616-3887 Email: jwallace@u.washington.edu Clones are derived from the human BAC library RPci-11. For BAC library availability, please contact Pieter de Jong (pietere@joeing.med.buffalo.edu). Clones may be purchased from BACPAC Resources (http://bacpac.med.buffalo.edu/ordering_bac.htm) or from Resear h Genetics (info@resgen.com). BAC end Web Server: http://www.htsc.washington.edu Plate: 1116 row: C column: 6 Seq primer: SP6 Class: BAC ends High quality sequence stop: 851. Location/Qualifiers 1..851 /organism="Homo sapiens" /db_xref="taxon:9606" /clone="Plate=1116 Col=6 Row=C" /clone_id="RPci-11 Human Male BAC Library" /sex="male" /note="Vector: pBACe3.6; Site_1: EcoRI; Site_2: EcoRI; Male blood DNA was isolated from one randomly chosen donor and partially digested with a combination of EcoRI and EcoRII Methyase. Size selected DNA was cloned into the pBACe3.6 vector at EcoRI sites"		
BASE COUNT	220 a	188 c	197 g
ORIGIN			245 t 1 others
Query Match	50.0%; Score 48;	DB 17;	Length 851;
Best Local Similarity	79.5%;	Pred. No. 4.9e-05;	
Matches 70;	Conservative 0;	Mismatches 15;	Indels 3;
			Gaps 1;
OY	12	CTCCGAGCGAAGTAATAATTATACACCACCTTTACACTAGACTCTTTGT---	AGAA 68
Db	156	CACCGAGCGAAGTAATAATTATACCAATTAATTCCTGACCTTTCTGTAGAGA	97
OY	69	AAGCCAATGAGTAGTGCACTCAATAGT 96	
Db	96	AAGCCAATGAGTAGTGCAATACCTTATGT 69	
RESULT 15			
LOCUS	AQ000365/c		
DEFINITION	CIT-HSP.2283L7.TR CIT-HSP Homo sapiens genomic clone 2283L7, DNA sequence.	632 bp	DNA linear GSS 26-JUN-1998

```

ACCESSION      A0000365
VERSION        A0000365.1
KEYWORDS       GSS.
SOURCE         human.
ORGANISM       Homo sapiens
               Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
               Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE      1 (bases 1 to 632)
               Adams,M.D., Rounsley,S.D., Zhao,S., Field,C.E., Bass,S., Linher,K.,
               Golden,K., Berry,K., Granger,D., Suh,E., Wible,C., Shizuya,H.,
               Simon,M. and Venter,J.C.
               Use of a random BAC End Sequence Database for Sequence-Ready Map
               Building (1998)
JOURNAL        Unpublished (1998)
COMMENT        Other-GSSs: CIT-HSP-2283L7.TF
               Contact: Mark Adams
               Department of Eukaryotic Genomics
               The Institute for Genomic Research
               9712 Medical Center Dr., Rockville, MD 20850, USA
               Tel.: 301 838 0200
               Fax: 301 838 0208
               Email: mdadams@tigr.org
               Clones are available from Research Genetics (info@resgen.com). BAC
               end search page:
               http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html
               Seq primer: M13 Reverse
               Class: BAC ends.
FEATURES       Location/Qualifiers
               source          1..632
                               /organism="Homo sapiens"
                               /db_xref="GDB:7148074"
                               /db_xref="taxon:9606"
                               /clone="2283L7"
                               /clone_id="CIT-HSP"
                               /sex="Male"
                               /cell_type="Sperm"
                               /note="Vector: pbeloBAC11; Site_1: HindIII; Site_2:
                               HindIII"
BASE COUNT    154 a      143 c      151 g      184 t
ORIGIN
Query Match   49.6%; Score 47.6; DB 17; Length 632;
Best Local Similarity 76.6%; Pred. No. 6e-05;
Matches 72; Conservative 0; Mismatches 19; Indels 3; Gaps 1;
QY      6 CTGACGACTCCTGAGGAGATATTAATTATACACCATCTTACAGCTAGACCTCTTTGTA 65
      11 | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      155 CTTGCCCCCTGAGGAGATATTAATTATTAAGCCCTCTTACAACTAGATCTTTCTGTA 96
      11 | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY      66 GAA---AAGGCAATGAGAGTGAAGTCCATAAGT 96
      11 | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      95 AACGAGAGGATAAATGAGATGAAGTCCCTTATGT 62
      11 | | | | | | | | | | | | | | | | | | | | | | | | | | | |

```


GenCore version 5.1.5
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 2, 2003, 11:30:30 ; Search time 38.4 Seconds

(without alignments)
766.692 Million cell updates/sec

Title: US-09-719-554-3_COPY_3065_3160

Perfect score: 96
Sequence: 1 gccgcctgagctccctcagag.....tgagtgagtgagtcataagt 96

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 441362 seqs, 15338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : Issued_Patents_NA:*
1: /cgn2_6/plodata/2/ina/5A.COMB.seq:*
2: /cgn2_6/plodata/2/ina/5B.COMB.seq:*
3: /cgn2_6/plodata/2/ina/6A.COMB.seq:*
4: /cgn2_6/plodata/2/ina/6B.COMB.seq:*
5: /cgn2_6/plodata/2/ina/PCITUS.COMB.seq:*
6: /cgn2_6/plodata/2/ina/backfilest1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	44.8	46.7	80246	US-09-078-294-4	Sequence 4, Appli
2	44.8	46.7	80595	US-09-078-294-3	Sequence 3, Appli
3	39.4	41.0	2503	US-08-691-563C-60	Sequence 60, Appli
4	28.4	29.6	434	US-09-149-476-30	Sequence 30, Appli
5	27	28.1	9060	US-08-378-313-20	Sequence 20, Appli
6	26.6	27.7	3635	US-08-252-626A-1	Sequence 1, Appli
7	25.4	26.5	483	US-08-286-819A-24	Sequence 24, Appli
8	25.4	26.5	483	US-08-980-357-24	Sequence 24, Appli
9	25.4	26.5	10851	US-08-286-819A-16	Sequence 16, Appli
10	25.4	26.5	10851	US-08-980-357-16	Sequence 16, Appli
11	25.2	26.2	874	US-08-858-207A-43	Sequence 43, Appli
12	25.2	26.2	1250	US-08-117-083-59	Sequence 59, Appli
13	25.2	26.2	1791	US-09-157-257-7	Sequence 15, Appli
14	24.8	25.8	857	US-09-268-364-15	Sequence 5, Appli
15	24.6	25.6	243	US-08-755-587-5	Sequence 5, Appli
16	24.6	25.6	2961	US-08-177-437-1	Sequence 1, Appli
17	24.4	25.4	864	US-08-204-675-3	Sequence 3, Appli
18	24.4	25.4	864	US-08-660-754-3	Sequence 3, Appli
19	24.4	25.4	864	US-08-796-364-3	Sequence 3, Appli
20	24.4	25.4	864	US-08-796-364-3	Sequence 3, Appli
21	24.4	25.4	4933	PCT-US95-02520-3	Sequence 2, Appli
22	24.4	25.4	4933	US-08-204-675-2	Sequence 2, Appli
23	24.4	25.4	4933	US-08-660-754-2	Sequence 2, Appli
24	24.4	25.4	4933	US-08-796-364-2	Sequence 2, Appli
25	24.4	25.4	5394	PCT-US95-02520-2	Sequence 2, Appli
26	24.4	25.4	5394	US-08-688-376-1	Sequence 1, Appli
27	24.2	25.2	211	US-09-813-133A-3	Sequence 3, Appli
				US-07-593-657-12	Sequence 12, Appli

28	24.2	25.2	7452	3	US-08-592-500-1	Sequence 1, Appli
29	24.2	25.2	7452	3	US-08-195-006-1	Sequence 1, Appli
30	24.2	25.2	7452	5	PCT-US94-07644A-1	Sequence 1, Appli
31	24	25.0	156	1	US-07-593-657-9	Sequence 9, Appli
32	24	25.0	2061	2	US-08-835-170-1	Sequence 1, Appli
33	24	25.0	2061	4	US-09-359-257-1	Sequence 1, Appli
34	24	25.0	2061	4	US-09-371-674-1	Sequence 1, Appli
35	24	25.0	2327	2	US-08-835-170-3	Sequence 3, Appli
36	24	25.0	2327	4	US-09-359-257-3	Sequence 3, Appli
37	24	25.0	2327	4	US-09-371-674-3	Sequence 3, Appli
38	24	25.0	2796	4	US-09-221-017B-693	Sequence 693, App
39	24	25.0	4342	4	US-09-338-907-107	Sequence 107, App
40	24	25.0	4342	4	US-09-218-207-107	Sequence 107, App
41	24	25.0	4582	4	US-09-338-907-118	Sequence 118, App
42	24	25.0	4582	4	US-09-218-207-118	Sequence 118, App
43	24	25.0	4686	4	US-09-338-907-117	Sequence 117, App
44	24	25.0	4686	4	US-09-218-207-117	Sequence 117, App
45	24	25.0	4875	4	US-09-338-907-114	Sequence 114, App

ALIGNMENTS

RESULT 1
US-09-078-294-4
; Sequence 4, Application US/09078294
; Patent No. 6265211
; GENERAL INFORMATION:
; APPLICANT: Choo, Kong-Hong Andy
; APPLICANT: Du Sart, Desiree
; APPLICANT: Cancilla, Michael R.
; TITLE OF INVENTION: A NOVEL NUCLEIC ACID MOLECULE
; FILE REFERENCE: Davies Col
; CURRENT APPLICATION NUMBER: US/09/078, 294
; CURRENT FILING DATE: 1998-05-13
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 4
; LENGTH: 80246
; TYPE: DNA
; ORGANISM: Nucleotide sequence of NC-contlg
US-09-078-294-4

Query Match 46.7%; Score 44.8; DB 4; Length 80246;
Best Local Similarity 77.3%; Pred. No. 2e-06; Mismatches 17; Indels 3; Gaps 1;
Matches 68; Conservative 0;

DB 12 CTCCTGAGGAGTAAATTAACACCATCTTACCTAGTACCTTTGTAGAA--- 68
DB 55967 CACCTGAGGAGTAAATTAACACCATCTTACCTAGTACCTTTGTAGAAAGG 56026

QY 69 AAGCAATGAGTGAAGTGCCTAAGT 96
DB 56027 AAGCAATGAGTGAAGTGCCTAAGT 56054

RESULT 2
US-09-078-294-3
; Sequence 3, Application US/09078294
; Patent No. 6265211
; GENERAL INFORMATION:
; APPLICANT: Choo, Kong-Hong Andy
; APPLICANT: Du Sart, Desiree
; APPLICANT: Cancilla, Michael R.
; TITLE OF INVENTION: A NOVEL NUCLEIC ACID MOLECULE
; FILE REFERENCE: Davies Col
; CURRENT APPLICATION NUMBER: US/09/078, 294
; CURRENT FILING DATE: 1998-05-13
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 3
; LENGTH: 80595
; TYPE: DNA

US-09-078-294-3

Query Match	46.7%;	Score	44.8;	DB	4;	Length	80595;
Best Local Similarity	77.3%;	Pred. No.	2e-06;				
Matches	68;	Conservative	0;	Mismatches	17;	Indels	3;
						Gaps	1

[illegible]

RESULT 3
US-08-691-563C-60

```

1 GENERAL INFORMATION:
2 APPLICANT: Herve PERRON
3 APPLICANT: Frederic BESEME
4 APPLICANT: Frederic BEDIN
5 APPLICANT: Glaucia PARANHOS-BACCALA
6 APPLICANT: Florence KOMURIAN-PRADEL
7 APPLICANT: Colette JOLIVET
8 APPLICANT: Bernard MANDRAND
9 TITLE OF INVENTION: VIRAL MATERIAL AND NUCLEOTIDE FRAGMENTS
10 TITLE OF INVENTION: ASSOCIATED WITH MULTIPLE SCLEROSIS, FOR DIAGNOSTIC, PROPHYLACTIC
11 TITLE OF INVENTION: THERAPEUTIC PURPOSES
12 NUMBER OF SEQUENCES: 92
13 CORRESPONDENCE ADDRESS:

```

```

? NAME: Berridge, William P.
? REGISTRATION NUMBER: 30.024
? REFERENCE/DOCKET NUMBER: WPB 38588
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: 703-836-6400
? TELEFAX: 703-836-2787
? INFORMATION FOR SEQ ID NO: 60:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 2503 base pairs
? TYPE: nucleotide
? STRANDEDNESS: single
? TOPOLOGY: linear
? MOLECULE TYPE: CDNA
? OS-08-691-563C-60

```

	Query Match	41.0%;	Score 39.4;	DB 3;	Length 2503;
	Best Local Similarity	76.6%;	Pred. No. 5;ge-05;		
	Matches 59;	Conservative 2;	Mismatches 13;	Indels 3;	Gaps
OY	12 CTCTGAGGGAAATATAATTATACACCATCTTACAGCTAACCCTTTTGTAAGA--- 68				
Dd	1496 CACCTGAGGGAAACACAATTAATTCACATATATATCTCAGCGTATCTTTCTGTATAGAGGG 1555				
OY	69 AAGCGAAATGCACTCAA 85				

Db 1556 AAGCAATGGAGTGA 1572

```

RESULT 4
US-09-149-476-30
Sequence 30, Application US/091149476
Patent No. 6420526
GENERAL INFORMATION:
APPLICANT: Rosen et al.
TITLE OF INVENTION: 166 Human Secreted proteins
FILE REFERENCE: P200201
CURRENT APPLICATION NUMBER: US/09/149,476
CURRENT FILING DATE: 1998-09-08
EARLIER APPLICATION NUMBER: PCT/US98/04493
EARLIER FILING DATE: 1998-03-06
EARLIER APPLICATION NUMBER: 60/040,162
EARLIER FILING DATE: 1997-03-07
EARLIER APPLICATION NUMBER: 60/040,333
EARLIER FILING DATE: 1997-03-07
EARLIER APPLICATION NUMBER: 60/038,621
EARLIER FILING DATE: 1997-03-07
EARLIER APPLICATION NUMBER: 60/040,626
EARLIER FILING DATE: 1997-03-07
EARLIER APPLICATION NUMBER: 60/040,334
EARLIER FILING DATE: 1997-03-07
EARLIER APPLICATION NUMBER: 60/040,336
EARLIER FILING DATE: 1997-03-07
EARLIER APPLICATION NUMBER: 60/040,163
EARLIER FILING DATE: 1997-03-07
EARLIER APPLICATION NUMBER: 60/047,600
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,615
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,597
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,502
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,633
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,583
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,617
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,618
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,503
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,592
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,581
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,584
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,500
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,587
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,492
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,598
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,613
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,582
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,596
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,612
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,632
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,601
EARLIER FILING DATE: 1997-05-23

```


[illegible]

EARLIER	FILING DATE:	1997-08-22
EARLIER	APPLICATION NUMBER:	60/047,595
EARLIER	FILING DATE:	1997-05-23
EARLIER	APPLICATION NUMBER:	60/047,599
EARLIER	FILING DATE:	1997-05-23
EARLIER	APPLICATION NUMBER:	60/047,588
EARLIER	FILING DATE:	1997-05-23
EARLIER	APPLICATION NUMBER:	60/047,585
EARLIER	FILING DATE:	1997-05-23
EARLIER	APPLICATION NUMBER:	60/047,586
EARLIER	FILING DATE:	1997-05-23
EARLIER	APPLICATION NUMBER:	60/047,590
EARLIER	FILING DATE:	1997-05-23
EARLIER	APPLICATION NUMBER:	60/047,594
EARLIER	FILING DATE:	1997-05-23
EARLIER	APPLICATION NUMBER:	60/047,589
EARLIER	FILING DATE:	1997-05-23
EARLIER	APPLICATION NUMBER:	60/047,593
EARLIER	FILING DATE:	1997-05-23
EARLIER	APPLICATION NUMBER:	60/047,614
EARLIER	FILING DATE:	1997-05-23
EARLIER	APPLICATION NUMBER:	60/043,578
EARLIER	FILING DATE:	1997-04-11
EARLIER	APPLICATION NUMBER:	60/043,576
EARLIER	FILING DATE:	1997-04-11
EARLIER	APPLICATION NUMBER:	60/047,501
EARLIER	FILING DATE:	1997-05-23
EARLIER	APPLICATION NUMBER:	60/043,670
EARLIER	FILING DATE:	1997-04-11
EARLIER	APPLICATION NUMBER:	60/056,632
EARLIER	FILING DATE:	1997-08-22
EARLIER	APPLICATION NUMBER:	60/056,664
EARLIER	FILING DATE:	1997-08-22
EARLIER	APPLICATION NUMBER:	60/056,876
EARLIER	FILING DATE:	1997-08-22
EARLIER	APPLICATION NUMBER:	60/056,881
EARLIER	FILING DATE:	1997-08-22
EARLIER	APPLICATION NUMBER:	60/056,909
EARLIER	FILING DATE:	1997-08-22
EARLIER	APPLICATION NUMBER:	60/056,875
EARLIER	FILING DATE:	1997-08-22
EARLIER	APPLICATION NUMBER:	60/056,862
EARLIER	FILING DATE:	1997-08-22
EARLIER	APPLICATION NUMBER:	60/056,887
EARLIER	FILING DATE:	1997-08-22
EARLIER	APPLICATION NUMBER:	60/056,908
EARLIER	FILING DATE:	1997-08-22
EARLIER	APPLICATION NUMBER:	60/048,964
EARLIER	FILING DATE:	1997-06-06
EARLIER	APPLICATION NUMBER:	60/057,650
EARLIER	FILING DATE:	1997-09-05
EARLIER	APPLICATION NUMBER:	60/056,884
EARLIER	FILING DATE:	1997-08-22
EARLIER	APPLICATION NUMBER:	60/057,669
EARLIER	FILING DATE:	1997-09-05
EARLIER	APPLICATION NUMBER:	60/049,610
EARLIER	FILING DATE:	1997-06-13
EARLIER	APPLICATION NUMBER:	60/061,060
EARLIER	FILING DATE:	1997-10-02

Query Match	29.68;	Score 28.4;	DB 4;	Length 434;
Best Local Similarity	60.38;	Pred. No. 0.3;		
Matches 41;	Conservative 3;	Mismatches 24;	Indels 0;	Gaps 0;

QY 28 AAATATTACACCACCATTTACAGCTGAGACTCTTTTGTAGAAAAAGCGCAATGSAGTGAA GT 87
||||| : |||||
Db 190 AAATGTATAAYTAFTGCCATTC TGACATGTTATTTAGCAAARGA AAAAAGAGTAATTC 249

QY	88	GCCATAAG	95
Db	250	TACATCAG	257

RESULT 5
US-08-378-313-20/c
; Sequence 20, Application US/08378313
; Patent No. 6207881
; GENERAL INFORMATION:
; APPLICANT: THEOLOGIS, ATHANASIOS
; APPLICANT: SATO, TAKAHIDO
; TITLE OF INVENTION: CONTROL OF FRUIT RIPENING THROUGH
; TITLE OF INVENTION: GENETIC CONTROL OF ACC SYNTHASE SYNTHESIS
; NUMBER OF SEQUENCES: 34
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORRISON & FOERSTER
; STREET: 755 Page Mill Road
; CITY: Palo Alto
; STATE: California
; COUNTRY: USA
; ZIP: 94304-1018
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/378,313
; FILING DATE:
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/862,493
; FILING DATE: 02-Apr-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: MURASHIGE, KATE H.
; REGISTRATION NUMBER: 29,959
; REFERENCE/DOCKET NUMBER: 29190-20002.20
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 856-5600
; TELEFAX: (415) 494-0792
; TELEX: 706141
; INFORMATION FOR SEQ ID NO: 20:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 9060 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; FEATURE:
; NAME/KEY: CDS
; LOCATION: join(2704..2880, 2968..3099, 3183..3344, 3810
; LOCATION: ..4376, 4463..4903)
US-08-378-313-20

Query Match 28.1%; Score 27; DB 4; Length 9060;
Best Local Similarity 66.1%; Pred. No. 2.4;
Matches 39; Conservative 0; Mismatches 20; Indels 0; Gaps 0;

QY 19 GGGAGCTATATATATACACCATCTTACAGCTAGACCTCTTTGTGAAAAGCCAAAT 77
DB 7306 GGGAGCTATATATATATACACCATCTTACAGCTAGACCTCTTTGTGAAAAGCCAAAT 7248

RESULT 6
US-08-252-626A-1/c
; Sequence 1, Application US/08252626A.
; Patent No. 5385269
; GENERAL INFORMATION:
; APPLICANT: Earp, Henry S.
; APPLICANT: Graham, Douglas K.
; APPLICANT: Dawson, Thomas L.
; APPLICANT: Mullaney, David L.
; APPLICANT: Snodgrass, Hiram R.
; TITLE OF INVENTION: Isolated DNA Encoding C-MER
; TITLE OF INVENTION: Protooncogene
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Kenneth D. Sibley

STREET: P.O. Drawer 34009
; CITY: Charlotte
; STATE: No. 5585269th Carolina
; COUNTRY: USA
; ZIP: 28234
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/252,626A
; FILING DATE: 02-JUN-1994
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Sibley, Kenneth D.
; REGISTRATION NUMBER: 31,665
; REFERENCE/DOCKET NUMBER: 5470-81
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (919) 881-3140
; TELEFAX: (919) 881-3175
; TELEX: 575102
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3635 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 138..3137
US-08-252-626A-1

Query Match 27.7%; Score 26.6; DB 1; Length 3635;
Best Local Similarity 60.3%; Pred. No. 2.5;
Matches 44; Conservative 0; Mismatches 29; Indels 0; Gaps 0;

QY 24 GTATTAATTTATACACCATCTTACAGCTAGACCTCTTTGTGAAAAGCCAAATGAGTG 83
DB 3445 GTACAAAGCTTAAATATACAGGACGCTGAGCTTAAATTTAGATATGCAAGTAAAGTG 3386

QY 84 AAGTGCCATAAGT 96
DB 3385 AATGGAATAGT 3373

RESULT 7
US-08-286-819A-24/c
; Sequence 24, Application US/08286819A
; Patent No. 5871910
; GENERAL INFORMATION:
; APPLICANT: ARTHUR, MICHEL
; APPLICANT: DURTA-MALEN, SYLVIE
; APPLICANT: MOLINAS, CATHERINE
; APPLICANT: COURVALIN, PATRICE
; TITLE OF INVENTION: POLYPEPTIDES IMPLICATED IN THE
; TITLE OF INVENTION: EXPRESSION OF RESISTANCE TO GLYCOPROTEINS, IN PARTICULAR
; TITLE OF INVENTION: IN GRAM-POSITIVE BACTERIA, NUCLEOTIDE SEQUENCE CODING FOR
; THESE POLYPEPTIDES AND USE FOR DIAGNOSIS
; NUMBER OF SEQUENCES: 54
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
; ADDRESSEE: P.C.
; STREET: 1755 S. Jefferson Davis Highway, Suite 400
; CITY: Arlington
; STATE: Virginia
; COUNTRY: U.S.A.
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/286,819A
FILING DATE: 05-AUG-1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/174,682
FILING DATE: 28-DEC-1993
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/917,146
FILING DATE: 10-AUG-1992
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/FR/91/00855
FILING DATE: 29-OCT-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: FR 9013579
FILING DATE: 31-OCT-1990
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: OBLON, NO. 5871910man F.
REGISTRATION NUMBER: 24,618
REFERENCE/DOCKET NUMBER: 660-060-0 PCT
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 413-3000
TELEFAX: (703) 413-2220
TELEX: 248855 OPAT UR
INFORMATION FOR SEQ ID NO: 24:
SEQUENCE CHARACTERISTICS:
LENGTH: 483 base pairs
TYPE: nucleic acid
STRANDEDNESS: unknown
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 1..483
US-08-286-819A-24

Query Match 26.5%; Score 25.4; DB 2; Length 483;
Best Local Similarity 61.2%; Pred. No. 3.6;
Matches 41; Conservative 0; Mismatches 26; Indels 0; Gaps 0;

QY 21 GAAGTAAATTTATACACCATCTTACAGCTAGACCTCTTTGTAGAAAAAGCAATGGA 80
Db 188 GGATTAAGATTATACATTATCTATCATCTCTGAAATTCACAGTACGATTAATGGA 129

QY 81 GTGAGT 87
Db 128 GTCAAGT 122

RESULT 8
US-08-980-357-24/c
Sequence 24, Application US/08980357
Patent No. 6013508
GENERAL INFORMATION:
APPLICANT: ARTHUR, MICHEL
APPLICANT: DURTA-MALEN, SYLVIE
APPLICANT: MOLINAS, CATHERINE
APPLICANT: COUVALIN, PATRICE
TITLE OF INVENTION: POLYPEPTIDES IMPLICATED IN THE
TITLE OF INVENTION: EXPRESSION OF RESISTANCE TO GLYCOPOLYMERES, IN PARTICULAR
TITLE OF INVENTION: IN GRAM-POSITIVE BACTERIA, NUCLEOTIDE SEQUENCE CODING FOR
NUMBER OF SEQUENCES: 54
CORRESPONDENCE ADDRESS:
ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
ADDRESSEE: P.C.
STREET: 1755 S. Jefferson Davis Highway, Suite 400
CITY: Arlington
STATE: Virginia

COUNTRY: U.S.A.
ZIP: 22202
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/980,357
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/286,819
FILING DATE: 05-AUG-1994
APPLICATION NUMBER: US 08/174,682
FILING DATE: 28-DEC-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/917,146
FILING DATE: 10-AUG-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/FR/91/00855
FILING DATE: 29-OCT-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: FR 9013579
FILING DATE: 31-OCT-1990
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: OBLON, NO. 6013508man F.
REGISTRATION NUMBER: 24,618
REFERENCE/DOCKET NUMBER: 660-060-0 PCT
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 413-3000
TELEFAX: (703) 413-2220
TELEX: 248855 OPAT UR
INFORMATION FOR SEQ ID NO: 24:
SEQUENCE CHARACTERISTICS:
LENGTH: 483 base pairs
TYPE: nucleic acid
STRANDEDNESS: unknown
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 1..483
US-08-980-357-24

Query Match 26.5%; Score 25.4; DB 3; Length 483;
Best Local Similarity 61.2%; Pred. No. 3.6;
Matches 41; Conservative 0; Mismatches 26; Indels 0; Gaps 0;

QY 21 GAAGTAAATTTATACACCATCTTACAGCTAGACCTCTTTGTAGAAAAAGCAATGGA 80
Db 188 GGATTAAGATTATACATTATCTATCATCTCTGAAATTCACAGTACGATTAATGGA 129

QY 81 GTGAGT 87
Db 128 GTCAAGT 122

RESULT 9
US-08-286-819A-16/c
Sequence 16, Application US/08286819A
Patent No. 5871910
GENERAL INFORMATION:
APPLICANT: ARTHUR, MICHEL
APPLICANT: DURTA-MALEN, SYLVIE
APPLICANT: MOLINAS, CATHERINE
APPLICANT: COUVALIN, PATRICE
TITLE OF INVENTION: POLYPEPTIDES IMPLICATED IN THE
TITLE OF INVENTION: EXPRESSION OF RESISTANCE TO GLYCOPOLYMERES, IN PARTICULAR
TITLE OF INVENTION: IN GRAM-POSITIVE BACTERIA, NUCLEOTIDE SEQUENCE CODING FOR
NUMBER OF SEQUENCES: 54
CORRESPONDENCE ADDRESS:

ADDRESS: OBLON, SPIVAK, MCLELLAND, MAIER & NEUSTADT,
ADDRESSEE: P.C. S. Jefferson Davis Highway, Suite 400
CITY: Arlington
STATE: Virginia
COUNTRY: U.S.A.
ZIP: 22202
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/286,819A
FILING DATE: 05-AUG-1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/174,682
FILING DATE: 28-DEC-1993
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/917,146
FILING DATE: 10-AUG-1992
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/FR/91/00855
FILING DATE: 29-OCT-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: FR 9013579
FILING DATE: 31-OCT-1990
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Oblon, No. 5871910man F.
REGISTRATION NUMBER: 24,618
REFERENCE/DOCKET NUMBER: 660-060-0 PCT
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 413-3000
TELEFAX: (703) 413-2220
TELEX: 248855 OPAT UR
INFORMATION FOR SEQ ID NO: 16:
SEQUENCE CHARACTERISTICS:
LENGTH: 10851 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: unknown
MOLECULE TYPE: DNA (genomic)
US-08-286-819A-16
Query Match 26.5%; Score 25.4; DB 2; Length 10851;
Best Local Similarity 61.2%; Pred. No. 9.4;
Matches 41; Conservative 0; Mismatches 26; Indels 0; Gaps 0;
QY 21 GAAGTAAATTAATACACCATCTTACAGCTGCTTTGTAGAAAAGCAATGGA 80
DB 10303 GGAATTAAGATTATACATATCATCATCTCTGAAATTCACAGTAGCAATGGA 10244
QY 81 GTGAGT 87
DB 10243 GTCAAGT 10237
RESULT 10
US-08-980-357-16/c
; Sequence 16, Application US/08980357
; Patent No. 6013508
; GENERAL INFORMATION:
; APPLICANT: ARTHUR, MICHEL
; APPLICANT: DURKA-MALEN, SYLVIE
; APPLICANT: MOLINAS, CATHERINE
; APPLICANT: COURVALIN, PATRICE
; TITLE OF INVENTION: POLYPEPTIDES IMPLICATED IN THE
; TITLE OF INVENTION: EXPRESSION OF RESISTANCE TO GLYCOPROTEINS, IN PARTICULAR
; TITLE OF INVENTION: IN GRAM-POSITIVE BACTERIA, NUCLEOTIDE SEQUENCE CODING FOR

TITLE OF INVENTION: THESE POLYPEPTIDES AND USE FOR DIAGNOSIS
NUMBER OF SEQUENCES: 54
CORRESPONDENCE ADDRESS:
ADDRESSEE: OBLON, SPIVAK, MCLELLAND, MAIER & NEUSTADT,
ADDRESSEE: P.C.
CITY: Arlington
STATE: Virginia
COUNTRY: U.S.A.
ZIP: 22202
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/980,357
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/286,819
FILING DATE: 05-AUG-1994
APPLICATION NUMBER: US 08/174,682
FILING DATE: 28-DEC-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/917,146
FILING DATE: 10-AUG-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/FR/91/00855
FILING DATE: 29-OCT-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: FR 9013579
FILING DATE: 31-OCT-1990
ATTORNEY/AGENT INFORMATION:
NAME: Oblon, No. 6013508man F.
REGISTRATION NUMBER: 24,618
REFERENCE/DOCKET NUMBER: 660-060-0 PCT
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 413-3000
TELEFAX: (703) 413-2220
TELEX: 248855 OPAT UR
INFORMATION FOR SEQ ID NO: 16:
SEQUENCE CHARACTERISTICS:
LENGTH: 10851 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: unknown
MOLECULE TYPE: DNA (genomic)
US-08-980-357-16
Query Match 26.5%; Score 25.4; DB 3; Length 10851;
Best Local Similarity 61.2%; Pred. No. 9.4;
Matches 41; Conservative 0; Mismatches 26; Indels 0; Gaps 0;
QY 21 GAAGTAAATTAATACACCATCTTACAGCTGCTTTGTAGAAAAGCAATGGA 80
DB 10303 GGAATTAAGATTATACATATCATCATCTCTGAAATTCACAGTAGCAATGGA 10244
QY 81 GTGAGT 87
DB 10243 GTCAAGT 10237
RESULT 11
US-08-858-207A-43
; Sequence 43, Application US/08858207A
; Patent No. 6348328
; GENERAL INFORMATION:
; APPLICANT: Black, Michael
; APPLICANT: Hodgson, John
; APPLICANT: Knowles, David
; APPLICANT: Nicholas, Richard
; APPLICANT: Stodola, Robert

```

1 TITLE OF INVENTION: No. 63483286L Compounds
2 NUMBER OF SEQUENCES: 552
3 CORRESPONDENCE ADDRESS:
4 ADDRESSEE: Smithsonian Beecham Corporation
5 STREET: 709 Swedeland Road
6 CITY: King of Prussia
7 STATE: PA
8 COUNTRY: USA
9 ZIP: 19406-0939
10 COMPUTER READABLE FORM:
11 MEDIUM TYPE: Diskette
12 OPERATING SYSTEM: DOS
13 SOFTWARE: FastSeq for Windows Version 2.0
14 CURRENT APPLICATION DATA:
15 APPLICATION NUMBER: US//08/858, 207A
16 FILING DATE: 09-MAY-1997
17 CLASSIFICATION: 435
18 PRIOR APPLICATION DATA:
19 APPLICATION NUMBER: 60/017670
20 FILING DATE: 14-MAY-1996
21 ATTORNEY/AGENT INFORMATION:
22 NAME: Gimmil, Edward R.
23 REGISTRATION NUMBER: 38,891
24 REFERENCE/DOCKET NUMBER: P50475
25 TELECOMMUNICATION INFORMATION:
26 TELEPHONE: 610-270-4478
27 TELEFAX: 610-270-5090
28 TELEX:
29 INFORMATION FOR SEQ ID NO: 43:
30 SEQUENCE CHARACTERISTICS:
31 LENGTH: 874 base pairs
32 TYPE: nucleic acid
33 STRANDEDNESS: single
34 TOPOLOGY: linear
35 US-08-858-207A-43
36
37 Query Match          26.2%, Score 25.2; DB 4; Length 874;
38 Best Local Similarity 57.7%; Pred.No. 5.1;
39 Matches 45; Conservative 0; Mismatches 33; Indels 0; Gaps 0;
40
41 Oy      19 GGGAAGTAAATTAATTACACCCTTTGACGTACAACCCITTTGGTAGAAAAGCCAAATG 78
42         ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
43 Db     218 GAGAAGGATAAAGTGATCACCACCGTCGACCACTTAAGGCCTTACTGTAATGGCTACT 277
44           ||   ||   |   ||| ||
45 Qy      79 GAGTGAAGTCCCATAGT 96
46             ||   ||   |   ||| ||
47 Db     278 GATGACCTTAGAATAGT 295
48
49 RESULT 12
50 US-08-117-083-59
51 Sequence 59, Application US/08117083
52 Patent No. 5719054
53 GENERAL INFORMATION:
54 APPLICANT: Boursnell, Michael E.
55 APPLICANT: Inglis, Stephen C.
56 APPLICANT: Munro, Alan J.
57 TITLE OF INVENTION: Recombinant Virus Vectors Encoding Human
58                 Papilloma Virus Proteins
59 NUMBER OF SEQUENCES: 70
60 CORRESPONDENCE ADDRESS:
61 ADDRESSEE: Walter H. Dreger
62 STREET: 4 Embarcadero Center, Suite 3400
63 CITY: San Francisco
64 STATE: CA
65 COUNTRY: USA
66 ZIP: 94111
67 COMPUTER READABLE FORM:
68 MEDIUM TYPE: Floppy disk
69 COMPUTER: IBM PC compatible
70 OPERATING SYSTEM: PC-DOS/M5-DOS
71 SOFTWARE: Patentin Release #1.0, Version #1.25

```

```

: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/117,083
: FILING DATE: 10-SEP-1993
: CLASSIFICATION: 435
: ATTORNEY/AGENT INFORMATION:
: NAME: Dreger, Walter H.
: REGISTRATION NUMBER: 24,190
: REFERENCE/DOCKET NUMBER: A-58783
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 415-781-1989
: TELEFAX: 415-398-3249
: TELEX: 910 277299
: INFORMATION FOR SEQ ID NO: 59:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 1250 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: double
: TOPOLOGY: linear
: MOLECULE TYPE: cDNA
US-08-117-083-59

Query Match          26.2%; Score 25.2; DB 1; Length 1250;
Best Local Similarity 62.9%; Pred. No. 5.7;
Matches 39; Conservative 0; Mismatches 23; Indels 0; Gaps 0;

QY 22 AAGTATAATTATTAACACCATCTTACAGCTACCTCTTTGTAGAAAAAGCAATGGAG 81
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1186 AAAATATAATTATTAATATACCTTCTACAGCGCGCCATCATGAATTAACAGACGATGAT 1245

QY 82 TG 83
   ||
DB 1246 TG 1247

RESULT 13
US-09-157-257-7
: Sequence 7, Application US/09157257
: Patent No. 6375954
: GENERAL INFORMATION:
: APPLICANT: DUTTA, Sukanta K.
: APPLICANT: BISWAS, Biswajit
: APPLICANT: VEDUPALLI, Ramesh
: TITLE OF INVENTION: A SIZE-VARIABLE STRAIN-SPECIFIC PROTECTIVE ANTIGEN FOR
: FILE REFERENCE: 8172-9016
: CURRENT APPLICATION NUMBER: US/09/157,257
: CURRENT FILING DATE: 1998-09-18
: EARLIER APPLICATION NUMBER: 60/059,252
: EARLIER FILING DATE: 1997-09-18
: NUMBER OF SEQ ID NOS: 48
: SOFTWARE: PatentIn Ver. 2.0
: SEQ ID NO 7
: LENGTH: 1791
: TYPE: DNA
: ORGANISM: Ehrlichia risticii
: FEATURE:
: NAME/KEY: CDS
: LOCATION: (175)..(1677)
US-09-157-257-7

Query Match          26.2%; Score 25.2; DB 4; Length 1791;
Best Local Similarity 60.0%; Pred. No. 6.4;
Matches 42; Conservative 0; Mismatches 28; Indels 0; Gaps 0;

QY 19 GGGAGTATAATTATTAACACCATCTTACAGCTTAACACCTCTTTGTAGAAAAAGCAATG 78
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 436 GGTAAAGATATATTAAGACATACACTACAGATAGACACTGTAAGTTTAAAGACCTGATA 555

QY 79 GAGTGAAGTG 88
   | ||| |
DB 556 GGAAGTAGCG 565

```

RESULT 14
US-09-268-364-15
; Sequence 15, Application US/09268364A
; Patent No. 6204063
; GENERAL INFORMATION:
; APPLICANT: Allen, Steve
; APPLICANT: Lee, Jian Ming
; APPLICANT: Lightner, Jonathan
; TITLE OF INVENTION: PLANT GLYCOLYSIS AND RESPIRATION ENZYMES
; FILE REFERENCE: BB-1154
; CURRENT APPLICATION NUMBER: US/09/268,364A
; EARLIER FILING DATE: 1999-03-15
; EARLIER APPLICATION NUMBER: 60/079,387
; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 15
; LENGTH: 857
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (11)
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (115)
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (744)
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (767)
; FEATURE:
; NAME/KEY: (794)..(795)
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (806)
; US-09-268-364-15

Query Match 25.8%; Score 24.8; DB 4; Length 857;
Best Local Similarity 72.7%; Pred. No. 7.1;
Matches 32; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

QY 40 CATCTACAGCTGACCTCTTTGTAGAAAAGCAATGAGATG 83
Db 295 CATCTACACCTGACCTATTTGCTTACAAAGCATGTGTAGAG 338

RESULT 15
US-08-755-587-5/C
; Sequence 5, Application US/08755587
; Patent No. 6045997
; GENERAL INFORMATION:
; APPLICANT: Futreal, Phillip A
; APPLICANT: Wooster, Richard F
; APPLICANT: Asmworth, Alan
; APPLICANT: Stratton, Michael R
; TITLE OF INVENTION: Materials and methods relating to the
; TITLE OF INVENTION: Identification and sequencing of the BRCA2 cancer
; TITLE OF INVENTION: susceptibility gene and uses thereof.
; NUMBER OF SEQUENCES: 222
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Bell Seltzer Park & Gibson
; STREET: 310 UCB Plaza, 3605 Glenwood Avenue, PO Drawer 31107
; CITY: Raleigh
; STATE: NC
; COUNTRY: USA
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/755,587
; FILING DATE: 25-NOV-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9523959.6
; FILING DATE: 23-NOV-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9525555.0
; FILING DATE: 14-DEC-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9617961.9
; FILING DATE: 28-AUG-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Kenneth D Sibley
; REGISTRATION NUMBER: 31,665
; REFERENCE/DOCKET NUMBER: 5405-135
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 243 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; US-08-755-587-5

Query Match 25.6%; Score 24.6; DB 3; Length 243;
Best Local Similarity 59.2%; Pred. No. 5.7;
Matches 42; Conservative 0; Mismatches 29; Indels 0; Gaps 0;

QY 7 TGGCAGCTCCTGAGGAGATATAATTATACACCATCTTACAGCTAGACCTCTTTGTAG 66
Db 119 TGGATCACCTGAGTGACAGAAATATATATACATACCTATAGAGGAGAGAAATATAAT 60

QY 67 AAAAGCAAT 77
Db 59 AAAAGCTAAT 49

Search completed: May 2, 2003, 12:49:12
Job time : 78.4 secs

GenCore version 5.1.5
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 2, 2003, 12:25:30 ; Search time 63.8897 Seconds
(without alignments)
1775.493 Million cell updates/sec

Title: US-09-719-554-3_COPY_3065_3160

Perfect score: 96

Sequence: 1 gccgcctgcacctcctgag.....tgagtgtaagtcacataagt 96

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 746064 seqs, 590810554 residues

Total number of hits satisfying chosen parameters: 1492128

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database :

Published_Applications_NA:*

- 1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
- 2: /cgn2_6/ptodata/1/pubpna/PC1_NEW_PUB.seq:*
- 3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
- 4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
- 5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
- 6: /cgn2_6/ptodata/1/pubpna/PC1S_PUBCOMB.seq:*
- 7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
- 8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
- 9: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
- 10: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
- 11: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
- 12: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*
- 13: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
- 14: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	67.4	70.2	764	7	US-08-979-847-126
2	67.4	70.2	800	7	US-08-979-847-127
3	64.6	67.3	1511	7	US-08-979-847-117
4	64	66.7	1359	9	US-10-125-237-2
5	64	66.7	1359	9	US-10-105-891-2
6	46.4	48.3	570	10	US-09-864-761-12831
7	44.8	46.7	584	10	US-09-924-400-316
8	44.8	46.7	800	10	US-09-924-400-307
9	44.8	46.7	800	10	US-09-825-301-11
10	44.8	46.7	800	10	US-08-924-400-307
11	44.8	46.7	800	10	US-08-810-936-307
12	44.8	46.7	800	10	US-09-429-755-307
13	44.8	46.7	829	9	US-09-924-400-317
14	44.8	46.7	829	10	US-09-810-936-317
15	42.4	44.2	502	10	US-09-783-590-4203
16	39.4	41.0	2503	7	US-08-979-847-56
17	31.4	32.7	42999	10	US-09-740-029-3
18	31	32.3	565	10	US-09-864-761-16360
19	30.6	31.9	487	9	US-09-918-995-6263

C 20	30	31.2	1691139	9	US-10-067-514-1	Sequence 1, Appl
C 21	29.4	30.6	5251	9	US-10-091-504-1431	Sequence 1431, Ap
C 22	29.4	30.6	5251	10	US-09-764-869-1431	Sequence 1431, Ap
C 23	28.4	28.6	434	9	US-09-809-391-30	Sequence 30, Appl
C 24	27.8	28.0	590	9	US-10-091-572-50	Sequence 50, Appl
C 25	27.8	28.0	590	9	US-09-764-891-686	Sequence 686, Appl
C 26	27.8	29.0	600	9	US-10-091-572-543	Sequence 543, Appl
C 27	27.8	29.0	600	9	US-09-764-891-686	Sequence 686, Appl
C 28	27.6	28.8	558	10	US-09-864-761-16507	Sequence 16507, A
C 29	27.4	28.5	2732	10	US-09-925-302-178	Sequence 178, Appl
C 30	27	28.1	423	10	US-09-783-590-10539	Sequence 10539, A
C 31	26.6	27.7	512	10	US-09-880-107-1443	Sequence 1443, Ap
C 32	26.6	27.7	2000	9	US-09-938-842A-4348	Sequence 4348, Ap
C 33	26.6	27.7	3608	9	US-10-174-590-433	Sequence 433, Appl
C 34	26.6	27.7	3608	9	US-10-176-758-433	Sequence 433, Appl
C 35	26.6	27.7	3608	9	US-10-175-737-433	Sequence 433, Appl
C 36	26.6	27.7	3608	9	US-10-173-706-433	Sequence 433, Appl
C 37	26.6	27.7	3608	9	US-10-175-738-433	Sequence 433, Appl
C 38	26.6	27.7	3608	9	US-10-175-752-433	Sequence 433, Appl
C 39	26.6	27.7	3608	9	US-10-176-482-433	Sequence 433, Appl
C 40	26.6	27.7	3608	9	US-10-176-757-433	Sequence 433, Appl
C 41	26.6	27.7	3608	9	US-10-176-913-433	Sequence 433, Appl
C 42	26.6	27.7	3608	9	US-10-180-552-433	Sequence 433, Appl
C 43	26.6	27.7	3608	9	US-10-180-557-433	Sequence 433, Appl
C 44	26.6	27.7	3608	9	US-10-173-700-433	Sequence 433, Appl
C 45	26.6	27.7	3608	9	US-10-174-572-433	Sequence 433, Appl

ALIGNMENTS

RESULT 1
US-08-979-847-126
Sequence 126, Application US/08979847
Publication No. US20030039664A1
GENERAL INFORMATION:
APPLICANT: PERRON, HERVE
APPLICANT: BESEME, FREDERIC
APPLICANT: BRDIN, FREDERIC
APPLICANT: PARANOS-BACCALA, GLAUCIA
APPLICANT: KOMURIAN-PRADEL, FLORENCE
APPLICANT: JOLIVET-REYNAUD, COLETTE
APPLICANT: MANDRAND, BERNARD
APPLICANT: GARSON, JEREMY
APPLICANT: TUBE, PHILIP
TITLE OF INVENTION: VIRAL MATERIAL AND NUCLEOTIDE FRAGMENTS
TITLE OF INVENTION: ASSOCIATED WITH MULTIPLE SCLEROSIS, FOR DIAGNOSTIC, PROPHY
NUMBER OF INVENTION: THERAPEUTIC PURPOSES
NUMBER OF SEQUENCES: 210
CORRESPONDENCE ADDRESS:
ADDRESSER: OLIVE & BERRIDGE, PLC
STREET: P.O. BOX 19928
CITY: ALEXANDRIA
STATE: VA
COUNTRY: USA
ZIP: 22320
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/979,847
FILING DATE: 26-NOV-1997
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: BERRIDGE, WILLIAM P.
REGISTRATION NUMBER: 30,024
REFERENCE/DOCKET NUMBER: WPP 39046A
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-836-6400
TELEFAX: 703-836-2787
INFORMATION FOR SEQ ID NO: 126:

SEQUENCE CHARACTERISTICS:
LENGTH: 764 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
US-08-979-847-126

Query Match 70.2%; Score 67.4; DB 7; Length 764;
Best Local Similarity 95.3%; Pred. No. 1.4e-13;
Matches 81; Conservative 0; Mismatches 1; Indels 3; Gaps 1;

QY 12 CTCCTGAGGAGATATAATTAATTAACACCATCTTACAGCTAGACCTCTTTGTAGAAA-- 69
|||||
Db 549 CTCCTGAGGAGATATAATTAATTAACACCATCTTACAGCTAGACCTCTTTGTAGAAAAG 608
QY 70 -AGGCAATGAGTGAAGTGCACATA 93
|||||
Db 609 GAGGCAATGAGTGAAGTGCACATA 633

RESULT 2

US-08-979-847-127
Sequence 127, Application US/08979847
Publication No. US20030039664A1
GENERAL INFORMATION:
APPLICANT: PERRON, HERVE
APPLICANT: BESEME, FREDERIC
APPLICANT: BEDIN, FREDERIC
APPLICANT: PARANHOS-BACCALA, GLAUCIA
APPLICANT: KOMURIAN-PRADEL, FLORENCE
APPLICANT: JOLIVET-REYNAUD, COLETTE
APPLICANT: MANDRAND, BERNARD
APPLICANT: GARSON, JEREMY
APPLICANT: TUKE, PHILIP
TITLE OF INVENTION: VIRAL MATERIAL AND NUCLEOTIDE FRAGMENTS
TITLE OF INVENTION: ASSOCIATED WITH MULTIPLE SCLEROSIS, FOR DIAGNOSTIC, PROPHYLACTIC
TITLE OF INVENTION: THERAPEUTIC PURPOSES
NUMBER OF SEQUENCES: 210
CORRESPONDENCE ADDRESS:
ADDRESSEE: OLIFF & BERRIDGE, PLC
STREET: P.O. BOX 19928
CITY: ALEXANDRIA
STATE: VA
COUNTRY: USA
ZIP: 22320
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
FILING DATE: 26-NOV-1997
APPLICATION NUMBER: US/08/979,847
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: BERRIDGE, WILLIAM P.
REGISTRATION/DOCKET NUMBER: 30,024
REFERENCE/DOCKET NUMBER: WPB 39046A
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-836-6400
TELEFAX: 703-836-2787
INFORMATION FOR SEQ ID NO: 127:
SEQUENCE CHARACTERISTICS:
LENGTH: 800 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
US-08-979-847-127

Query Match 70.2%; Score 67.4; DB 7; Length 800;
Best Local Similarity 95.3%; Pred. No. 1.5e-13;

Matches 81; Conservative 0; Mismatches 1; Indels 3; Gaps 1;

QY 12 CTCCTGAGGAGATATAATTAATTAACACCATCTTACAGCTAGACCTCTTTGTAGAAA-- 69
|||||
Db 551 CTCCTGAGGAGATATAATTAATTAACACCATCTTACAGCTAGACCTCTTTGTAGAAAAG 610
QY 70 -AGGCAATGAGTGAAGTGCACATA 93
|||||
Db 611 GAGGCAATGAGTGAAGTGCACATA 635

RESULT 3

US-08-979-847-117
Sequence 117, Application US/08979847
Publication No. US20030039664A1
GENERAL INFORMATION:
APPLICANT: PERRON, HERVE
APPLICANT: BESEME, FREDERIC
APPLICANT: BEDIN, FREDERIC
APPLICANT: PARANHOS-BACCALA, GLAUCIA
APPLICANT: KOMURIAN-PRADEL, FLORENCE
APPLICANT: JOLIVET-REYNAUD, COLETTE
APPLICANT: MANDRAND, BERNARD
APPLICANT: GARSON, JEREMY
APPLICANT: TUKE, PHILIP
TITLE OF INVENTION: VIRAL MATERIAL AND NUCLEOTIDE FRAGMENTS
TITLE OF INVENTION: ASSOCIATED WITH MULTIPLE SCLEROSIS, FOR DIAGNOSTIC, PROPHYLACTIC
TITLE OF INVENTION: THERAPEUTIC PURPOSES
NUMBER OF SEQUENCES: 210
CORRESPONDENCE ADDRESS:
ADDRESSEE: OLIFF & BERRIDGE, PLC
STREET: P.O. BOX 19928
CITY: ALEXANDRIA
STATE: VA
COUNTRY: USA
ZIP: 22320
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
FILING DATE: 26-NOV-1997
APPLICATION NUMBER: US/08/979,847
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: BERRIDGE, WILLIAM P.
REGISTRATION/DOCKET NUMBER: 30,024
REFERENCE/DOCKET NUMBER: WPB 39046A
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-836-6400
TELEFAX: 703-836-2787
INFORMATION FOR SEQ ID NO: 117:
SEQUENCE CHARACTERISTICS:
LENGTH: 1511 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
US-08-979-847-117

Query Match 67.3%; Score 64.6; DB 7; Length 1511;
Best Local Similarity 92.0%; Pred. No. 1.7e-12;
Matches 80; Conservative 0; Mismatches 4; Indels 3; Gaps 1;

QY 13 TCCTGAGGAGATATAATTAATTAACACCATCTTACAGCTAGACCTCTTTGTAGAAA--A 69
|||||
Db 124 TCCTGAGGAGATATAATTAATTAACACCATCTTACAGCTAGACCTCTTTGTAGAAAAG 183
QY 70 AGGCAATGAGTGAAGTGCACATA 96
|||||
Db 184 GAGGCAATGAGTGAAGTGCACATA 210


```
RESULT 4
US-10-125-237-2
; Sequence 2, Application US/10125237
; Publication No. US20030022329A1
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Zhou, Ping
; APPLICANT: Asundi, Vinod
; APPLICANT: Ren, Feiyan
; APPLICANT: Zhao, Qiong A.
; APPLICANT: Xue, Aidong J.
; APPLICANT: Zhang, Jie
; APPLICANT: Wehrman, Tom
; APPLICANT: Wang, Jian-Rui
; TITLE OF INVENTION: No. US20030022329A1el Nucleic Acids and
; FILE REFERENCE: 791CIP2ADIV
; CURRENT APPLICATION NUMBER: US/10/125,237
; PRIOR FILING DATE: 2002-04-17
; PRIOR APPLICATION NUMBER: 09/668,317
; PRIOR FILING DATE: 2000-09-22
; PRIOR APPLICATION NUMBER: 09/552,929
; NUMBER OF SEQ ID NOS: 91
; SOFTWARE: pt_fl_genes Version 2.0
; SEQ ID NO 2
; LENGTH: 1359
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (114)..(1280)
US-10-125-237-2

Query Match          66.7%; Score 64; DB 9; Length 1359;
Best Local Similarity 90.9%; Pred. No. 2,6e-12;
Matches 80; Conservative 0; Mismatches 5; Indels 3; Gaps 1;

QY 12 CTCCTGAGGAGATATAATTAATTAACACCATCTTACAGCTAGACCTCTTTGTAGAA--- 68
      ||||| || ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 307 CTCCTAGAGAGATATAATTAATTAATTAACACCATCTTACAGCTAGACCTCTTTGTAGAAAGG 366
QY 69 AAGGCAATGGAGTGAAGTGCATATGT 96
      | ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 367 AAGGCAATGGAGTGAAGTGCATATGT 394

RESULT 5
US-10-105-891-2
; Sequence 2, Application US/10105891
; Publication No. US20030073099A1
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Zhou, Ping
; APPLICANT: Asundi, Vinod
; APPLICANT: Ren, Feiyan
; APPLICANT: Zhao, Qiong A.
; APPLICANT: Xue, Aidong J.
; APPLICANT: Zhang, Jie
; APPLICANT: Wehrman, Tom
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Dimanac, Radoje T.
; TITLE OF INVENTION: No. US20030073099A1el Nucleic Acids and
; FILE REFERENCE: 791CIP2A
; CURRENT APPLICATION NUMBER: US/10/105,891
; PRIOR FILING DATE: 2002-03-25
; PRIOR APPLICATION NUMBER: 09/668,317
; PRIOR FILING DATE: 2000-09-22
```

```
;; PRIOR APPLICATION NUMBER: 09/552,929
;; PRIOR FILING DATE: 2000-04-18
;; NUMBER OF SEQ ID NOS: 91
;; SOFTWARE: pt_fl_genes Version 2.0
;; SEQ ID NO 2
;; LENGTH: 1359
;; TYPE: DNA
;; ORGANISM: Homo sapiens
;; FEATURE:
;; NAME/KEY: CDS
;; LOCATION: (114)..(1280)
US-10-105-891-2

Query Match          66.7%; Score 64; DB 9; Length 1359;
Best Local Similarity 90.9%; Pred. No. 2,6e-12;
Matches 80; Conservative 0; Mismatches 5; Indels 3; Gaps 1;

QY 12 CTCCTGAGGAGATATAATTAATTAACACCATCTTACAGCTAGACCTCTTTGTAGAA--- 68
      ||||| || ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 307 CTCCTAGAGAGATATAATTAATTAATTAACACCATCTTACAGCTAGACCTCTTTGTAGAAAGG 366
QY 69 AAGGCAATGGAGTGAAGTGCATATGT 96
      | ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 367 AAGGCAATGGAGTGAAGTGCATATGT 394

RESULT 6
US-09-864-761-12831/c
; Sequence 12831, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL F
; FILE REFERENCE: Aecm1ca-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; PRIOR FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
```

```
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 12831
; LENGTH: 570
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC011778.3
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 2.1
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.4
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.2
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 0.93
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.2
US-09-864-761-12831
```

```
Query Match          48.3%; Score 46.4; DB 10; Length 570;
Best Local Similarity 78.4%; Pred. No. 2.2e-06;
Matches 69; Conservative 0; Mismatches 16; Indels 3; Gaps 1;
```

```
OY 12 CTCCTGAGGAGATTAATTAACACCATCTTACAGCTAGACCTCTTTGT---AGAA 68
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 453 CACCTGAGGAGATTAATTAACATCTCTGAGCTTACCTTTCTGTGAAGAG 394
```

```
OY 69 AAGGCAATGAGTGAAGTCCCATAGT 96
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 393 AAGGCAATGAGTGAATACATTAGT 366
```

RESULT 7

```
US-09-924-400-316/c
; Sequence 316, Application US/09924400
; Patent No. US20020165371A1
; GENERAL INFORMATION:
; APPLICANT: Fridakis, Tony N.
; APPLICANT: Reed, Steven G.
; APPLICANT: Smith, John M.
; APPLICANT: Misher, Lynda E.
; APPLICANT: Dillon, Davin C.
; APPLICANT: Retter, Marc W.
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A. W.
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Day, Craig H.
; APPLICANT: Li, Samuel X.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF BREAST CANCER
; FILE REFERENCE: 210121.419C12
; CURRENT APPLICATION NUMBER: US/09/924,400
; CURRENT FILING DATE: 2001-08-07
; NUMBER OF SEQ ID NOS: 340
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 316
; LENGTH: 584
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-924-400-316
```

```
Query Match          46.7%; Score 44.8; DB 9; Length 584;
Best Local Similarity 77.3%; Pred. No. 8e-06;
Matches 68; Conservative 0; Mismatches 17; Indels 3; Gaps 1;
```

```
OY 12 CTCCTGAGGAGATTAATTAACACCATCTTACAGCTAGACCTCTTTGTAGAA--- 68
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 339 CACCTGAGGAGATTAACATCTCTGAGCTTACCTTTCTGTGAAGAGG 280
```

```
OY 69 AAGGCAATGAGTGAAGTCCCATAGT 96
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 279 AAGGCAATGAGTGAATACATTAGT 252
```

RESULT 8

```
US-09-810-936-316/c
; Sequence 316, Application US/09810936
; Patent No. US20020068285A1
; GENERAL INFORMATION:
; APPLICANT: Fridakis, Tony N.
; APPLICANT: Reed, Steven G.
; APPLICANT: Smith, John M.
; APPLICANT: Misher, Lynda E.
; APPLICANT: Dillon, Davin C.
; APPLICANT: Retter, Marc W.
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A. W.
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Day, Craig H.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE
; TITLE OF INVENTION: THERAPY AND DIAGNOSIS OF BREAST CANCER
; FILE REFERENCE: 210121.419C11
; CURRENT APPLICATION NUMBER: US/09/810,936
; CURRENT FILING DATE: 2001-03-16
; NUMBER OF SEQ ID NOS: 334
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 316
; LENGTH: 584
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-810-936-316
```

```
Query Match          46.7%; Score 44.8; DB 10; Length 584;
Best Local Similarity 77.3%; Pred. No. 8e-06;
Matches 68; Conservative 0; Mismatches 17; Indels 3; Gaps 1;
```

```
OY 12 CTCCTGAGGAGATTAATTAACACCATCTTACAGCTAGACCTCTTTGTAGAA--- 68
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 339 CACCTGAGGAGATTAACATCTCTGAGCTTACCTTTCTGTGAAGAGG 280
```

```
OY 69 AAGGCAATGAGTGAAGTCCCATAGT 96
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 279 AAGGCAATGAGTGAATACATTAGT 252
```

RESULT 9

```
US-09-924-400-307
; Sequence 307, Application US/09924400
; Patent No. US20020165371A1
; GENERAL INFORMATION:
; APPLICANT: Fridakis, Tony N.
; APPLICANT: Reed, Steven G.
; APPLICANT: Smith, John M.
; APPLICANT: Misher, Lynda E.
; APPLICANT: Dillon, Davin C.
; APPLICANT: Retter, Marc W.
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A. W.
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Day, Craig H.
; APPLICANT: Li, Samuel X.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF BREAST CANCER
; FILE REFERENCE: 210121.419C12
; CURRENT APPLICATION NUMBER: US/09/924,400
; CURRENT FILING DATE: 2001-08-07
; NUMBER OF SEQ ID NOS: 340
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 307
; LENGTH: 800
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-924-400-307
```



```

; TITLE OF INVENTION: AND DIAGNOSIS OF BREAST CANCER
; FILE REFERENCE: 210121.419C12
; CURRENT APPLICATION NUMBER: US/09/924,400
; CURRENT FILING DATE: 2001-08-07
; NUMBER OF SEQ ID NOS: 340
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 317
; LENGTH: 829
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-924-400-317

Query Match          46.7%; Score 44.8; DB 9; Length 829;
Best Local Similarity 77.3%; Pred. No. 9e-06;
Matches 68; Conservative 0; Mismatches 17; Indels 3; Gaps 1;

QY 12 CTCCTGAGGAATGATTAATATATACACCATCTTACAGCTAGACCTTTGTAGAA--- 68
DB 582 CACCTGAGGGAATGAGATTACATATCTCTGAGCTTGTGTAAGAGGG 641
QY 69 AAGGCAATGAGTGAAGTGCATTAAGT 96
DB 642 AAGGCAATGAGTGAATACCTTAAGT 669

RESULT 14
US-09-810-936-317
; Sequence 317, Application US/09810936
; Patent No. US20020068285A1
; GENERAL INFORMATION:
; APPLICANT: Frudakis, Tony N.
; APPLICANT: Reed, Steven G.
; APPLICANT: Smith, John M.
; APPLICANT: Misher, Linda E.
; APPLICANT: Dillon, Davin C.
; APPLICANT: Retter, Marc W.
; APPLICANT: Wang, Aljun
; APPLICANT: Skelky, Yasir A.W.
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Day, Craig H.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE
; TITLE OF INVENTION: THERAPY AND DIAGNOSIS OF BREAST CANCER
; FILE REFERENCE: 210121.419C11
; CURRENT APPLICATION NUMBER: US/09/810,936
; CURRENT FILING DATE: 2001-03-16
; NUMBER OF SEQ ID NOS: 334
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 317
; LENGTH: 829
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-810-936-317

Query Match          46.7%; Score 44.8; DB 10; Length 829;
Best Local Similarity 77.3%; Pred. No. 9e-06;
Matches 68; Conservative 0; Mismatches 17; Indels 3; Gaps 1;

QY 12 CTCCTGAGGGAATTAATATATACACCATCTTACAGCTAGACCTTTGTAGAA--- 68
DB 582 CACCTGAGGGAATGAGATTACATATCTCTGAGCTTGTGTAAGAGGG 641
QY 69 AAGGCAATGAGTGAAGTGCATTAAGT 96
DB 642 AAGGCAATGAGTGAATACCTTAAGT 669

RESULT 15
US-09-783-590-4203/c
; Sequence 4203, Application US/09783590
; Patent No. US20020110850A1
; GENERAL INFORMATION:
; APPLICANT: Dillion, Patrick J.
; APPLICANT: Haseltine, William A.
```

```

; APPLICANT: Li, Haodong
; APPLICANT: Rosen, Craig A.
; APPLICANT: Ruben, Steven M.
; TITLE OF INVENTION: Human Genes, Sequences, and Expression Products 16.2
; FILE REFERENCE: PO-16.2c1
; CURRENT APPLICATION NUMBER: US/09/783,590
; CURRENT FILING DATE: 2000-02-15
; PRIOR APPLICATION NUMBER: 08/420,856
; PRIOR FILING DATE: 1995-04-12
; PRIOR APPLICATION NUMBER: 08/346,731
; PRIOR FILING DATE: 1994-11-21
; NUMBER OF SEQ ID NOS: 12485
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4203
; LENGTH: 502
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (20)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (121)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (140)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (144)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (168)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (192)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (330)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (355)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (371)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (372)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (374)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (387)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (395)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (409)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (412)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (413)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (438)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (439)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
```

```

LOCATION: (449)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc feature
LOCATION: (452)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc feature
LOCATION: (460)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc feature
LOCATION: (462)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc feature
LOCATION: (463)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc feature
LOCATION: (471)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc feature
LOCATION: (472)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc feature
LOCATION: (474)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc feature
LOCATION: (477)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc feature
LOCATION: (480)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc feature
LOCATION: (481)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc feature
LOCATION: (482)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc feature
LOCATION: (484)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc feature
LOCATION: (489)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc feature
LOCATION: (490)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc feature
LOCATION: (491)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc feature
LOCATION: (494)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc feature
LOCATION: (495)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc feature
LOCATION: (502)
OTHER INFORMATION: n equals a,t,g, or c
US-09-783-590-4203

```

```

Query Match 44.2%; Score 42.4; DB 10; Length 502;
Best Local Similarity 81.7%; Pred. No. 5,le-05;
Matches 49; Conservative 0; Mismatches 11; Indels 0; Gaps 0;
QY 12 CTCCTGAGGAGTAATTAATTAACACCACTTACAGCTAGACCTCTTTGTAGAAAG 71
Db 83 CACCTGAGGAGTAATTAACACCACTTACAGCTAGACCTCTTTGTAGAAAG 24

```

Search completed: May 2, 2003, 14:43:14
 Job time : 66.8897 secs

GenCore version 5.1.5
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 2, 2003, 11:21:20 ; Search time 406.862 Seconds

(Without alignments)
6795.341 Million cell updates/sec

Title: US-09-719-554-3_COPY_6956_7050

Sequence: 1 ccgcctatcgccaaagtc.....accacacagcccaactca 95

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl:*

1: gb_ba:*

2: gb_htg:*

3: gb_in:*

4: gb_om:*

5: gb_ov:*

6: gb_pat:*

7: gb_ph:*

8: gb_pl:*

9: gb_pr:*

10: gb_ro:*

11: gb_sts:*

12: gb_sy:*

13: gb_un:*

14: gb_vi:*

15: em_ba:*

16: em_fun:*

17: em_hum:*

18: em_in:*

19: em_mu:*

20: em_om:*

21: em_or:*

22: em_ov:*

23: em_pat:*

24: em_ph:*

25: em_pl:*

26: em_ro:*

27: em_sts:*

28: em_un:*

29: em_vi:*

30: em_htg_hum:*

31: em_htg_in:*

32: em_htg_other:*

33: em_htg_mus:*

34: em_htg_pla:*

35: em_htg_rtd:*

36: em_htg_mam:*

37: em_htg_vtc:*

38: em_sy:*

39: em_htgo_hum:*

40: em_htgo_mus:*

41: em_htgo_other:*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	95	100.0	3372	6	AX000964	AX000964 Sequence
2	95	100.0	3372	6	AX027478	AX027478 Sequence
3	95	100.0	3372	9	AF072502	AF072502 Homo sapi
4	95	100.0	10499	6	AX007980	AX007980 Sequence
5	95	100.0	56093	6	AX329572	AX329572 Sequence
6	95	100.0	56093	9	HSAC000064	AC000064 Human BAC
7	95	100.0	149194	9	AC007566	AC007566 Homo sapi
8	94	98.9	2930	6	AC355872	AC355872 Sequence
9	94	98.9	2930	9	AF208161	AF208161 Homo sapi
10	94	98.9	2946	6	AR177269	AR177269 Sequence
11	92.6	97.5	7582	6	AX000966	AX000966 Sequence
12	92.6	97.5	7582	6	AX027480	AX027480 Sequence
13	90.4	95.2	374	6	BD003152	BD003152 Secretary
14	90.4	95.2	374	6	180071	180071 Sequence 47
15	87.8	92.4	2599	6	AX007978	AX007978 Sequence
16	85.4	89.9	2006	6	AX000959	AX000959 Sequence
17	85.4	89.9	2006	6	AX027473	AX027473 Sequence
18	85.4	89.9	2006	9	AF072503	AF072503 Homo sapi
19	82.2	86.5	170754	2	AC012403	AC012403 Homo sapi
20	82.2	86.5	177866	2	AL354698	AL354698 Homo sapi
21	82.2	86.5	180954	9	AL160281	AL160281 Human DNA
22	79.2	83.4	134369	9	AC104825	AC104825 Homo sapi
23	79.2	83.4	176249	2	AC069497	AC069497 Homo sapi
24	77.6	81.7	70352	9	AL592310	AL592310 Human DNA
25	77.6	81.7	73435	2	AC091015	AC091015 Homo sapi
26	77.6	81.7	155021	9	HS774610	AL034410 Human DNA
27	77.6	81.7	159851	9	AC036114	AC036114 Homo sapi
28	77.6	81.7	163631	9	AC009276	AC009276 Homo sapi
29	77.6	81.7	165260	9	AC024341	AC024341 Homo sapi
30	77.6	81.7	168177	2	AC023914	AC023914 Homo sapi
31	77.6	81.7	170919	9	AC078847	AC078847 Homo sapi
32	77.6	81.7	186822	2	AC073235	AC073235 Homo sapi
33	77.6	81.7	195344	2	AC025248	AC025248 Homo sapi
34	77.6	81.7	196039	2	AC055757	AC055757 Homo sapi
35	76	80.0	1740	6	AX080036	AX080036 Sequence
36	76	80.0	46575	2	AC080036	AC080036 Homo sapi
37	76	80.0	114621	9	AC005187	AC005187 Homo sapi
38	76	80.0	137947	9	HS4514	AL023581 Human DNA
39	76	80.0	153444	2	AC040948	AC040948 Homo sapi
40	76	80.0	173788	9	AC022555	AC022555 Homo sapi
41	76	80.0	182928	2	AP001161	AP001161 Homo sapi
42	76	80.0	183680	9	AC098859	AC098859 Homo sapi
43	76	80.0	186540	9	AC090341	AC090341 Homo sapi
44	76	80.0	187837	9	AC079065	AC079065 Homo sapi
45	76	80.0	205035	9	CNS00005	AL049870 Human Chr

ALIGNMENTS

RESULT 1

AX000964 3372 bp DNA linear PAT 10-MAR-2000

LOCUS

DEFINITION Sequence 9 from Patent WO9902696.

ACCESSION AX000964

VERSION AX000964.1 GI:7241206

KEYWORDS

SOURCE

ORGANISM

unidentified.

unidentified

unclassified.

REFERENCE 1 (bases 1 to 3372)

AUTHORS Beseme, F. and Blond, J.

TITLE ENDOGENETIC RETROVIRAL SEQUENCES, ASSOCIATED WITH AUTOIMMUNE DISEASES OR WITH PREGNANCY DISORDERS

JOURNAL Patent: WO 9902696-A 9 21-JAN-1999;

FEATURES BIO MERIEUX (FR); BESEME FREDERIC (FR)
source 1..3372
/organism="unidentified"
/db_xref="taxon:32644"
BASE COUNT 1047 a 835 c 711 g 779 t
ORIGIN

Query Match 100.0%; Score 95; DB 6; Length 3372;
Best Local Similarity 100.0%; Pred. No. 6.6e-23;
Matches 95; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCTGCTTATCGCCAAAGCTCTTCAGAGACAAGAGAGCCATTACCTGGAGAGA 60
|||||
DB 1887 CCTGCTTATCGCCAAAGCTCTTCAGAGACAAGAGAGCCATTACCTGGAGAGA 1946
CTGGCAACTGATTTTACCCACAAGCCCAAACTCA 95
|||||
DB 1947 CTGGCAACTGATTTTACCCACAAGCCCAAACTCA 1981
|||||

RESULT 2
AX027478 3372 bp DNA linear PAT 16-SEP-2000
LOCUS AX027478
DEFINITION Sequence 28 from Patent FR2788784.
ACCESSION AX027478
VERSION AX027478.1 GI:10188442
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
REFERENCE 1 (bases 1 to 3372)
AUTHORS Mallet,F., Voisset,C. and Paranhos,B.G.
JOURNAL Patient: FR 2788784-A 28-28-JUL-2000;
BIO MERIEUX (FR)

FEATURES
source 1..3372
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 1047 a 835 c 711 g 779 t
ORIGIN

Query Match 100.0%; Score 95; DB 6; Length 3372;
Best Local Similarity 100.0%; Pred. No. 6.6e-23;
Matches 95; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCTGCTTATCGCCAAAGCTCTTCAGAGACAAGAGAGCCATTACCTGGAGAGA 60
|||||
DB 1887 CCTGCTTATCGCCAAAGCTCTTCAGAGACAAGAGAGCCATTACCTGGAGAGA 1946
CTGGCAACTGATTTTACCCACAAGCCCAAACTCA 95
|||||
DB 1947 CTGGCAACTGATTTTACCCACAAGCCCAAACTCA 1981
|||||

RESULT 3
AF072502 3372 bp mRNA linear PRI 10-FEB-1999
LOCUS AF072502
DEFINITION Homo sapiens endogenous retrovirus W sequence.
ACCESSION AF072502
VERSION AF072502.1 GI:4262286
KEYWORDS
SOURCE Homo sapiens.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
REFERENCE 1 (bases 1 to 3372)
AUTHORS Blond,J.L., Beseme,F., Duret,L., Bouton,O., Bedin,F., Perron,H.,
Mandrand,B. and Mallet,F.
TITLE Molecular characterization and placental expression of HERV-W, a
JOURNAL new human endogenous retrovirus family
J. Virol. 73 (2), 1175-1185 (1999)

MEDLINE 99099005
PUBMED 9882319
REFERENCE 2 (bases 1 to 3372)
AUTHORS Blond,J.L., Beseme,F. and Mallet,F.
TITLE Direct Submission
JOURNAL Submitted (19-JUN-1998) UM103 CNRS-biomerieux, ENS Lyon, 46 allée
d'Italie, Lyon, Cedex 07 69364, France

FEATURES
source 1..3372
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="cl.p15r"
/issue="type="Placenta"
/clone.lib="Clontech 5'-stretch plus library, Cat number
HL5014a"
1..3372
/note="corresponds to pol"
repeat_region 1..3372
/rpl_family="HERV-W"
misc_feature 2785..2786
/note="putative splice junction"
3'UTR 3025..3351
/note="U3-R junction undetermined"
polyA_site 3350
BASE COUNT 1047 a 835 c 711 g 779 t
ORIGIN

Query Match 100.0%; Score 95; DB 9; Length 3372;
Best Local Similarity 100.0%; Pred. No. 6.6e-23;
Matches 95; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCTGCTTATCGCCAAAGCTCTTCAGAGACAAGAGAGCCATTACCTGGAGAGA 60
|||||
DB 1887 CCTGCTTATCGCCAAAGCTCTTCAGAGACAAGAGAGCCATTACCTGGAGAGA 1946
CTGGCAACTGATTTTACCCACAAGCCCAAACTCA 95
|||||
DB 1947 CTGGCAACTGATTTTACCCACAAGCCCAAACTCA 1981
|||||

RESULT 4
AX007980 10499 bp DNA linear PAT 06-SEP-2000
LOCUS AX007980
DEFINITION Sequence 3 from Patent WO9967395.
ACCESSION AX007980
VERSION AX007980.1 GI:9995677
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
REFERENCE 1 (bases 1 to 10499)
AUTHORS Perin,J.P., Rieger,F. and Alliel,P.M.
TITLE Nucleic sequence and deduced protein sequence family with human
endogenous retroviral motifs, and their uses
JOURNAL Patent: WO 9967395-A 3 29-DEC-1999;
INSR NAT SANTE RECH MED (FR); PERIN JEAN PIERRE (FR); RIEGER
FRANCOIS (FR); ALLIEL PATRICK M (FR)

FEATURES
source 1..10499
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 3048 a 2676 c 2280 g 2495 t
ORIGIN

Query Match 100.0%; Score 95; DB 6; Length 10499;
Best Local Similarity 100.0%; Pred. No. 8.1e-23;
Matches 95; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCTGCTTATCGCCAAAGCTCTTCAGAGACAAGAGAGCCATTACCTGGAGAGA 60
|||||
DB 6956 CCTGCTTATCGCCAAAGCTCTTCAGAGACAAGAGAGCCATTACCTGGAGAGA 7015
|||||


```

Oy 61 CTGGCACTGATTTTACCACAGCCCAACCTCA 95
      |||||||
Db 7016 CTGGCACTGATTTTACCACAGCCCAACCTCA 7050

RESULT 5
AX329572 56093 bp DNA linear PAT 09-JAN-2002
LOCUS Sequence 81 from Patent WO0194629.
DEFINITION AX329572
ACCESSION AX329572
VERSION AX329572.1 GI:18102550
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1
AUTHORS Young,P.E., Augustus,M., Carter,K.C., Ebner,R., Endress,G.,
            Horriqan,S., Soppet,D.R. and Weaver,Z.
            Cancer gene determination and therapeutic screening using signature
            gene sets
            Patent: WO 0194629-A 81 13-DEC-2001;
            Avalon Pharmaceuticals (US)
FEATURES
            Location/Qualifiers
            source          1..56093
                           /organism="Homo sapiens"
                           /db_xref="taxon:9606"
BASE COUNT 16164 a 12346 c 10702 g 16881 t
ORIGIN
Query Match 100.0%; Score 95; DB 6; Length 56093;
Best Local Similarity 100.0%; Pred. No. 11e-22;
Matches 95; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 CCTGCTTATGCGCAGCTCCTTCAGGAGAACAAAGAGCCATTACCTCGAGAGAGA 60
Db 34956 CCTGCTTATGCGCAGCTCCTTCAGGAGAACAAAGAGCCATTACCTCGAGAGAGA 35015

Oy 61 CTGGCACTGATTTTACCACAGCCCAACCTCA 95
      |||||||
Db 35016 CTGGCACTGATTTTACCACAGCCCAACCTCA 35050

RESULT 6
HSAC000064 56093 bp DNA linear PRI 13-NOV-1996
LOCUS Human BAC clone RG083M05 from 7q21-7q22, complete sequence.
DEFINITION HSAC000064
ACCESSION AC000064
VERSION AC000064.1 GI:1669369
KEYWORDS HTG.
SOURCE Homo sapiens.
            Homo sapiens.
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1 (bases 1 to 56093)
AUTHORS Pauley,A.
TITLE The sequence of H. sapiens BAC clone RG083M05
JOURNAL Unpublished (1996)
REFERENCE 2 (bases 1 to 56093)
AUTHORS Waterston,R.
TITLE Direct Submission
JOURNAL Submitted (13-NOV-1996)
COMMENT Genome Sequencing Center
            Department of Genetics, Washington University
            St. Louis, MO 63108, USA
            e-mail: saplens@watson.wustl.edu

```

NOTICE: This sequence may not represent the entire insert of this clone. It may be shorter because we only sequence overlapping sections once, or longer because we provide a small overlap between neighboring submissions.

This sequence was finished as follows unless otherwise noted:

all regions were double stranded or sequenced with an alternate chemistry; an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

SOURCE INFORMATION:

This clone is from the first release of the human BAC library. The library contains cloned DNA from a human male fibroblast cell line 978SK. For references see: Shizuya et al., Proc. Natl. Acad. Sci. 89:8794-8797 (1992); Kim et al., Genomics 34:213-218 (1996).

VECTOR: pBELO
Selection: chloramphenicol

NEIGHBORING SEQUENCE INFORMATION:

The orientation of this clone is unknown. Actual start of this clone is at base position 1 of H.RG083M05; actual end is at 56093 of H.RG083M05

This clone contains STS SW851725.

FEATURES

```

source          Location/Qualifiers
            1..56093
               /organism="Homo sapiens"
               /db_xref="taxon:9606"
               /chromosome="7"
               /map="7q21-7q22"
               /clone="H.RG083M05"
               /clone_11b="C17B-978SK-B"
               complement(838..1131)
               /rpl_family="ALU"
               <1360..16971
               /gene="WUGSC:H.RG083M05.1"
               join(<1360..1503,4181..4370,4587..4774,6422..6556,
               9483..9547,11631..11773,11864..12021,13131..13296,
               14885..14988,16349..16546,16837..16971)
               /gene="WUGSC:H.RG083M05.1"
               /note="ATPase: strong similarity to peroxisome
               biosynthesis protein PAB1 (PID:g1172019); coded for by
               human CDNA C04279 (NID:91467530)"
               /codon_start=1
               /protein_id="AAB46346.1"
               /db_xref="GI:1669371"
               /translation="KRLNIQKTELEAFSEAVWMPQSVLLDDLLAGLPAVEHEH
               SPDAORCEITCNVTKNKIDCDINKKIDLDIOHAKETGCVFARDFTLVNRAIHSRL
               SROISITREKRYVLTITDQKALRGFLPALSLSVNLHNRDRCMPKIGLHVEYROIAD
               TIQLPAAKCLKKERYPELFANPLPRQGIILVGPFGTKTLGAVIARBSRNFISV
               KGPFLSKYIGASDAQVADIFIRQAAKPCILFDEFESIAPRGHDVTGTDRAVNO
               LI70LDGVEGAGVYLAATSRPDLIDPALRPGRLDKCVCPDPDYTTISLESKTG
               OMHSFLVSRLEITNLVSDSLADVDVLDQHAVSVTDSFTGADLKALYNNLEALHG
               MLKMSRFLPDESKFNMYRIYFGSSYSESGNGTSSDLSGCLSAPESSMDNPGPI
               KKDOLFSPPLVLRATASQGCCEUVEORQDRLADISITKGYRSGSGEDSMNPGPI
               KRLAISQSHMLTALGHTRPISSEDMKFAEL"
               complement(4948..5130)
               /rpl_family="ALU"
               complement(4948..5130)
               /rpl_family="ALU"
               complement(6581..7133)
               /rpl_family="11"
               complement(7767..8037)
               /rpl_family="ALU"
               complement(8186..8472)
               /rpl_family="ALU"
               8473..8625
               /gene="WUGSC:H.RG083M05.1"
               /note="match to human 3' EST H75782 (NID:g1049794), bases
               287-444"
               8841..9161
               /gene="WUGSC:H.RG083M05.1"
               /note="match to human 5' EST H75921 (NID:g1050050), bases
               21-348"
               9481..9547
               /gene="WUGSC:H.RG083M05.1"
               /note="match to human 5' EST N22627 (NID:g1130501), bases
               276-343"
               complement(12612..12907)
               repeat_region
               misc_feature
               misc_feature
               misc_feature
               repeat_region

```

```

misc_feature      /rpt_family="ALU"
                  13670..13793
/gene="WUGSC:H.RG083M05.1"
/feature="match to human 5' EST H41382 (NID:917434), bases
143-266"
repeat_region    13794..13877
/rpt_family="ALU"
misc_feature      13878..13906
/gene="WUGSC:H.RG083M05.1"
/feature="match to human 5' EST H41382 (NID:917434), bases
30-58"
repeat_region    13907..14104
/rpt_family="ALU"
repeat_region    complement(14110..14137)
/rpt_family="L1"
repeat_region    complement(15618..15907)
/rpt_family="ALU"
repeat_region    17227..17322
/rpt_family="ALU"
misc_feature      18667..19235
/feature="match to human fetal brain 5' EST D61494
(NID:970409), bases 1-255, and to human 3' EST R07476
(NID:9759399)"
repeat_region    19550..19670
/rpt_family="ALU"
misc_feature      21507..37303
/feature="similarity to various SS-RNA virus polypeptides;
pseudogene: region of matches and close matches to
multiple human ESTs, see R68740 (NID:9842257)"
37316..37489
/feature="Grail prediction, score = 80"
exon             /evidence="not_experimental"
repeat_region    complement(38938..39224)
/rpt_family="ALU"
misc_feature      39225..39707
/feature="match to multiple human ESTs, see N30113
(NID:9114863)"
39800..40085
/rpt_family="ALU"
repeat_region    complement(40247..40538)
/rpt_family="ALU"
repeat_region    complement(40632..40924)
/rpt_family="ALU"
repeat_region    complement(42283..42891)
/rpt_family="ALU"
repeat_region    complement(45474..45613)
/rpt_family="ALU"
misc_feature      complement(45614..45737)
/feature="match to human 3' EST H48898 (NID:9988738), bases
129-333"
misc_feature      complement(46107..47026)
/feature="match to multiple human ESTs, see N81064
(NID:91243765), H48897 (NID:9988737), and M78831
(NID:9273146)"
repeat_region    complement(47027..47318)
/rpt_family="ALU"
misc_feature      complement(47365..47782)
/feature="match to multiple human ESTs, see W37495
(NID:9139089)"
47898..48115
/feature="match to human 5' EST H63306 (NID:91015138), bases
93-368"
repeat_region    complement(48116..48405)
/rpt_family="ALU"
misc_feature      complement(48406..48584)
/feature="match to human 3' EST N29952 (NID:91148472), bases
290-455, and 5' EST R42730 (NID:9765806)"
repeat_region    complement(48787..49405)
/rpt_family="ALU"
misc_feature      complement(49406..49534)
/feature="match to human 3' EST R65794 (NID:9838432), bases
309-440"
repeat_region    complement(49638..49672)

```

```

misc_feature      /rpt_family="ALU"
                  complement(49674..49890)
/feature="match to human 3' EST N29952 (NID:91148472) and 5'
EST N29938 (NID:91148458), sequences are from opposite
ends of the same clone"
gene             complement(49698..51806)
/feature="WUGSC:H.RG083M05.2"
CDS              complement(join(49698..49888,51575..51806))
/feature="WUGSC:H.RG083M05.2"
/feature="WUGSC:H.RG083M05.2"
/feature="coded for by human cDNAs W37389 (NID:91319205),
R65891 (NID:9838529), R65794 (NID:9838432) and R65794
(NID:9838432)"
/codon_start=1
/protein_id="AAB46345.1"
/db_xref="GI:1669370"
/translation="MPEYFPGCGILIRCFPGVYVIGDVYSVIDEDGKPYAQIRGF
IDQYCKSAALTLWLIPTLSSPRDQDPASIIIGPEDDLPKMEYLFVCHAPSEYFK
SRSPPTVPTRPKGYIWTWGVPTATIRKSVANHL"
exon             complement(51576..51758)
/feature="WUGSC:H.RG083M05.2"
repeat_region    complement(51576..51758)
/feature="Grail prediction, score = 86"
/feature="not_experimental"
misc_feature      complement(52052..52329)
/rpt_family="L1"
/feature="match to human EST M79192 (NID:9273505) base 2-289"
Query Match      100.0%: Score 95; DB 9; Length 56093;
Best Local Similarity 100.0%: Pred. No. 11e-22;
Matches 95; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY              1 CCGCCCTTATCGCCAGAGCTCTTCAGAGAAACAGAGCCATTACCTGAGAGAGA 60
Db              34956 CCGGCCCTTATCGCCAGAGCTCTTCAGAGAAACAGAGCCATTACCTGAGAGAGA 35015
QY              61 CTGGCAACTGATTTTACCCCAAGCCCAACCTCA 95
Db              35016 CTGGCAACTGATTTTACCCCAAGCCCAACCTCA 35050
RESULT 7
AC007566/c      149194 bp      DNA      linear      PRI 01-MAR-2002
LOCUS          AC007566
DEFINITION     Homo sapiens BAC clone CTB-10G5 from 7q21-7q22, complete sequence.
ACCESSION     AC007566
VERSION       AC007566.2 GI:11181861
KEYWORDS      HTG.
SOURCE        Homo sapiens.
ORGANISM      Homo sapiens.
REFERENCE     1 (bases 1 to 149194)
AUTHORS      Sulston, J.E. and Waterston, R.
TITLE        Toward a complete human genome sequence
JOURNAL      Genome Res. 8 (11), 1097-1108 (1998)
MEDLINE      99063792
PUBMED       9847074
REFERENCE     2 (bases 1 to 149194)
AUTHORS      Du, Z.
TITLE        The sequence of Homo sapiens BAC clone CTB-10G5
JOURNAL      Unpublished (2001)
REFERENCE     3 (bases 1 to 149194)
AUTHORS      Waterston, R.H.
TITLE        Direct Submission
JOURNAL      Submitted (15-MAY-1999) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
MO 63108, USA
REFERENCE     4 (bases 1 to 149194)
AUTHORS      Waterston, R.
TITLE        Direct Submission
JOURNAL      Submitted (02-OCT-2000) Department of Genetics, Washington
University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
REFERENCE     5 (bases 1 to 149194)

```

AUTHORS Waterston, R.H.
TITLE Direct Submission
JOURNAL Submitted (16-NOV-2000) Genome Sequencing Center, Washington University School of Medicine, 4444 Forest Park Parkway, St. Louis, MO 63108, USA
REFERENCE 6 (bases 1 to 149194)
AUTHORS Waterston, R.H.
TITLE Direct Submission
JOURNAL Submitted (03-JAN-2002) Genome Sequencing Center, Washington University School of Medicine, 4444 Forest Park Parkway, St. Louis, MO 63108, USA
REFERENCE 7 (bases 1 to 149194)
AUTHORS Waterston, R.H.
TITLE Direct Submission
JOURNAL Submitted (06-FEB-2002) Genome Sequencing Center, Washington University School of Medicine, 4444 Forest Park Parkway, St. Louis, MO 63108, USA
REFERENCE 8 (bases 1 to 149194)
AUTHORS Waterston, R.
TITLE Direct Submission
JOURNAL Submitted (01-MAR-2002) Department of Genetics, Washington University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA On Nov 10, 2000 this sequence version replaced g1:4835815.
COMMENT -----
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: http://genome.wustl.edu/gsc
Contact: sapient@wustl.wustl.edu
Summary Statistics
Center project name: R_RG010605

NOTICE: This sequence may not represent the entire insert of this clone. It may be shorter, because we only sequence overlapping clone sections once, or longer because we provide a small overlap between neighboring data submissions.
This sequence was finished as follows unless otherwise noted:
all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality > 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.
MAPPING INFORMATION:
The sequence of this clone was established as part of a mapping and sequencing collaboration between the NHGRI Chromosome 7 Mapping Project (Eric D. Green, Director), John D. McPherson in the Department of Genetics (Washington University), and the Washington University Genome Sequencing Center. For additional information about the map position of this sequence, see
http://www.nhgri.nih.gov/DIR/CTB/CHR7, send
mailto:egreen@nhgri.nih.gov, or see http://genome.wustl.edu/gsc
SOURCE INFORMATION:
Clone CTB-10G5 is from the first release of the human BAC library CTB-978SK-B. The library contains cloned DNA from the male fibroblast cell line 978SK. See: Shizuya et al., Proc. Natl. Acad. Sci. USA 89:8794-7 (1992); U-C. Kim et al., Genomics 34:213-8 (1996). This clone is available from Research Genetics, Inc. (http://www.resgen.com).
VECTOR: pBelBAC11
Selection: chloramphenicol
NEIGHBORING SEQUENCE INFORMATION:
The clone sequenced to the left is RP5-91H5, 200 base pair overlap. Actual start of this clone is at base position 195 of CTB-10G5; actual end is at base position 150532 of CTB-10G5.
The clone CTB-10G5 contains the entire sequence of CTB-83M5.
Location/Qualifiers
1..149194
FEATURES
source

/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="7"
/map="7q21-7q22"
/clone="CTB-10G5"
/clone_id="CTB-978SK-B"
1..1634
/rpt_family="L1"
2248..2388
/note="match to EST BG752883 (NID:g14063536)"
2248..2387
/note="match to EST AA149693 (NID:g1720635) zn99d06.r1"
2248..2387
/note="match to EST AW579261 (NID:g7254310)"
2248..2387
/note="match to EST BG766882 (NID:g14077535)"
2248..2387
/note="match to EST BI160365 (NID:g14620366)"
2248..2374
/note="similar to Homo sapiens EST BF758865 (NID:g12106765)"
2248..2287
/note="match to EST AV686676 (NID:g10288539)"
2253..2387
/note="match to EST AU123510 (NID:g10948226)"
2344..2387
/note="match to EST AV686676 (NID:g10288539)"
2696..3066
/rpt_family="MALR"
3108..3392
/rpt_family="Alu"
3540..3628
/note="match to EST AA425526 (NID:g2106267) zw48b03.r1"
3540..3628
/note="similar to Mus musculus EST BB253526 (NID:g8946272)"
3542..3632
/note="match to EST AV686676 (NID:g10288539)"
3542..3628
/note="match to EST AA149693 (NID:g1720635) zn99d06.r1"
3542..3628
/note="match to EST AU123510 (NID:g10948226)"
3542..3628
/note="match to EST BG766882 (NID:g14077535)"
3542..3628
/note="match to EST BI160365 (NID:g14620366)"
3542..3609
/note="match to EST AW579261 (NID:g7254310)"
3562..3628
/note="match to EST BE272564 (NID:g9146913)"
3598..3628
/note="match to EST BE299708 (NID:g9183456)"
3714..3785
/note="match to EST BG260659 (NID:g12770475)"
3717..3785
/note="similar to Mus musculus EST BE994936 (NID:g10678674)"
3719..3785
/note="match to EST AA149693 (NID:g1720635) zn99d06.r1"
3719..3785
/note="match to EST AA425526 (NID:g2106267) zw48b03.r1"
3719..3785
/note="match to EST AU123510 (NID:g10948226)"
3719..3785
/note="match to EST BE272564 (NID:g9146913)"
3719..3785
/note="match to EST BE299708 (NID:g9183456)"
3719..3785
/note="match to EST BG766882 (NID:g14077535)"
3719..3785
/note="match to EST BI160365 (NID:g14620366)"
3719..3785
/note="similar to Mus musculus EST BB253526

misc_feature (NID:98946272)"
3860..3862
/note="match to EST AV686676 (NID:g10288539)"
repeat_region 4730..4827
/rpt_family="L2"
repeat_region 4834..4925
/rpt_family="L1"
misc_feature 5715..5860
/note="match to EST AA425526 (NID:g2106267) zw48b03.r1"
misc_feature 5715..5860
/note="match to EST AU123510 (NID:g10948226)"
misc_feature 5715..5860
/note="match to EST BE272564 (NID:g9146913)"
misc_feature 5715..5860
/note="match to EST BE29708 (NID:g9183456)"
misc_feature 5715..5860
/note="match to EST BG260659 (NID:g12770475)"
misc_feature 5715..5860
/note="match to EST BG766882 (NID:g14077535)"
misc_feature 5715..5860
/note="match to EST BI160365 (NID:g14620366)"
misc_feature 5715..5860
/note="similar to Mus musculus EST BE994936"
Query Match 100.0%; Score 95; DB 9; Length 149194;
Best Local Similarity 100.0%; Pred. No. 1.3e-22;
Matches 95; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CCGGCTTATCGCCAGACTCTCTTCAGAGAAAGAACAGGCGCATTTACCTGGAGAGA 60
|||||
DB 87048 CCGGCTTATCGCCAGACTCTCTTCAGAGAAAGAACAGGCGCATTTACCTGGAGAGA 86989
|||||
QY 61 CTGGCACTGATTTTACCACAGCCCAACCTCA 95
|||||
DB 86988 CTGGCACTGATTTTACCACAGCCCAACCTCA 86954
|||||
RESULT 8
AX355872 2930 bp DNA linear PAT 06-FEB-2002
LOCUS AX355872
DEFINITION Sequence 1 from Patent WO0204678.
ACCESSION AX355872
VERSION AX355872.1 GI:18620523
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1
Keith,J.C., McCoy,J.M. and M1,S.
Methods and compositions for diagnosing and treating pre eclampsia
and gestational trophoblast disorders
Patent: WO 0204678-A 1 17-JAN-2002;
JOURNAL GENETICS INSTITUTE, INC. (US)
Location/Qualifiers
1. 2930
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
930..2546
/note="unnamed protein product"
/codon_start=1
/protein_id="CAD23000.1"
/db_xref="GI:18620524"
CDS
/translation="MALPYHIFLFTVLPSTLTAPPCRCMTSSSPVOEFLMROR
GNIDAPSYRSLSKGPTPTFAHTHMPNCHSATLCHNANTHWGKMINPSCPGGLG
TVCWTFYQTGMSDGGVODAREKHVKVISQLRVHGTSSPYKGLDLSKLAETLT
HRLVSLFNTTLGLHEVSAQNPNTNCMLPLNFRVYSIPVEQNNSTELINTSV
LVGPLVSNLEITHNTSLNCVKFSNTYNSOCIRKRVYPTQIVCLPSGIFPVCSTSA
YKLNSSSSMKFLSLVPPNTIYTPDOLYNYVISPRKRKRPILPFVIGAGVLAGS
TGIGGTTSTORYKLSOELNDGMDERVAOVLTLDOQLNSLAAYVQNRALDLTAE
RGCTCLFLDEECYVYVNSGIYTERKREIRDRIORAEELRNTGPGGLSQMPWILP
RSGCTCLFLDEECYVYVNSGIYTERKREIRDRIORAEELRNTGPGGLSQMPWILP
FLGPLAAIILLFLFGPCIFNLNVNFEVSSRIEAVKIQMEPKMOSKTKIYRRPLDRPAS
RSDVNDIKTTPPEISAAOPLLRPNAGSS"

BASE COUNT 842 a 800 c 571 g 717 t
ORIGIN
Query Match 98.9%; Score 94; DB 6; Length 2930;
Best Local Similarity 100.0%; Pred. No. 1.5e-22;
Matches 94; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2 CTGCTTATCGCCAGACTCTCTTCAGAGAAAGAACAGGCGCATTTACCTGGAGAGAC 61
|||||
DB 8 CTGCTTATCGCCAGACTCTCTTCAGAGAAAGAACAGGCGCATTTACCTGGAGAGAC 67
|||||
QY 62 TGGCACTGATTTTACCACAGCCCAACCTCA 95
|||||
DB 68 TGGCACTGATTTTACCACAGCCCAACCTCA 101
|||||
RESULT 9
AF208161 2930 bp mRNA linear PRI 22-FEB-2000
LOCUS AF208161
DEFINITION Homo sapiens syncytin precursor, mRNA, complete cds.
ACCESSION AF208161
VERSION AF208161.1 GI:6760400
KEYWORDS
SOURCE Homo sapiens.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1 (bases 1 to 2930)
M1,S., Lee,X., Li,X., Veldman,G.M., Finnerty,H., Racie,L.,
Lavallee,E., Tang,X.Y., Edouard,P., Howes,S., Keith,J.C. Jr. and
McCoy,J.M.
Syncytin is a captive retroviral envelope protein involved in human
placental morphogenesis
Nature 403 (6771), 785-789 (2000)
JOURNAL Nature 403 (6771), 785-789 (2000)
MEDLINE 20155476
PUBMED 10693809
REFERENCE 2 (bases 1 to 2930)
Sha,M., Lee,X., Li,X., Veldman,G.M., Finnerty,H., Racie,L.,
Lavallee,E., Tang,X., Edouard,P., Howes,S., Keith,J.C. Jr. and
McCoy,J.M.
Direct Submision
Submitted (26-NOV-1999) Genetics Institute, 87 Cambridge Park
Drive, Cambridge, MA 02140, USA
FEATURES
source
1. 2930
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="AJ172-25"
/tissue_type="testis"
1. 2930
/rpt_family="Human endogenous retrovirus HERV-W"
1..929
930..2546
/note="envelope protein"
/product="syncytin precursor"
/codon_start=1
/protein_id="AAF28334.1"
/db_xref="GI:6760401"
/translation="MALPYHIFLFTVLPSTLTAPPCRCMTSSSPVOEFLMROR
GNIDAPSYRSLSKGPTPTFAHTHMPNCHSATLCHNANTHWGKMINPSCPGGLG
TVCWTFYQTGMSDGGVODAREKHVKVISQLRVHGTSSPYKGLDLSKLAETLT
HRLVSLFNTTLGLHEVSAQNPNTNCMLPLNFRVYSIPVEQNNSTELINTSV
LVGPLVSNLEITHNTSLNCVKFSNTYNSOCIRKRVYPTQIVCLPSGIFPVCSTSA
YKLNSSSSMKFLSLVPPNTIYTPDOLYNYVISPRKRKRPILPFVIGAGVLAGS
TGIGGTTSTORYKLSOELNDGMDERVAOVLTLDOQLNSLAAYVQNRALDLTAE
RGCTCLFLDEECYVYVNSGIYTERKREIRDRIORAEELRNTGPGGLSQMPWILP
RSGCTCLFLDEECYVYVNSGIYTERKREIRDRIORAEELRNTGPGGLSQMPWILP
FLGPLAAIILLFLFGPCIFNLNVNFEVSSRIEAVKIQMEPKMOSKTKIYRRPLDRPAS
RSDVNDIKTTPPEISAAOPLLRPNAGSS"
sig_peptide
mat_peptide
misc_feature
930..989
990..2543
/product="syncytin"
1869..1880
/note="furin; cleavage site"

misc_feature 2043..2120 /evidence=not_experimental
/note="Region: Immunosuppressive region"
misc_feature 2235..2357
/note="transmembrane-region site"
3'UTR 2547..2930
BASE COUNT 842 a 800 c 571 g 717 t
ORIGIN

Query Match 98.9%; Score 94; DB 9; Length 2930;
Best Local Similarity 100.0%; Pred. No. 1.5e-22;
Matches 94; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 CTGCTTATGCGCAAGCTCTTCAGAGAACAAAGAGCCATTACCTCGAGAAGAC 61
|||||
DB 8 CTGCTTATGCGCAAGCTCTTCAGAGAACAAAGAGCCATTACCTCGAGAAGAC 67
|||||

QY 62 TGGCAACTGATTTTACCACCAAGCCCAAACTCA 95
|||||
DB 68 TGGCAACTGATTTTACCACCAAGCCCAAACTCA 101
|||||

RESULT 10
LOCUS AR177269 2946 bp DNA linear PAT 17-DEC-2001
DEFINITION Sequence 3 from patent US 6312921.
ACCESSION AR177269
VERSION AR177269.1 GI:17919624
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 2946)
AUTHORS Jacoby,K., McCoy,J.M., Lavallie,E.R., Racie,L.A., Evans,C.,
Merberg,D., M.S. and Treacy,M.
TITLE Secreted proteins and polynucleotides encoding them
JOURNAL Patent: US 6312921-A 3 06-NOV-2001;
FEATURES
source location/Qualifiers
1..2946
BASE COUNT 858 a 801 c 570 g 717 t
ORIGIN

Query Match 98.9%; Score 94; DB 6; Length 2946;
Best Local Similarity 100.0%; Pred. No. 1.5e-22;
Matches 94; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 CTGCTTATGCGCAAGCTCTTCAGAGAACAAAGAGCCATTACCTCGAGAAGAC 61
|||||
DB 6 CTGCTTATGCGCAAGCTCTTCAGAGAACAAAGAGCCATTACCTCGAGAAGAC 65
|||||

QY 62 TGGCAACTGATTTTACCACCAAGCCCAAACTCA 95
|||||
DB 66 TGGCAACTGATTTTACCACCAAGCCCAAACTCA 99
|||||

RESULT 11
LOCUS AX000966 7582 bp DNA linear PAT 10-MAR-2000
DEFINITION Sequence 11 from Patent WO9902696.
ACCESSION AX000966
VERSION AX000966.1 GI:7241208
KEYWORDS
SOURCE unidentified.
ORGANISM unidentified.
REFERENCE 1 (bases 1 to 7582)
AUTHORS Beseme,F. and Blond,J.
TITLE ENDOGENETIC RETROVIRAL SEQUENCES, ASSOCIATED WITH AUTOIMMUNE
DISEASES OR WITH PREGNANCY DISORDERS
JOURNAL Patent: WO 9902696-A 11 21-JAN-1999;
FEATURES BIO MERIEUX (FR); BESEME FREDERIC (FR)
location/Qualifiers

source 1..7582
/organism="unidentified"
/db_xref="taxon:32644"
BASE COUNT 2156 a 1876 c 1538 g 1796 t 216 others
ORIGIN

Query Match 97.5%; Score 92.6; DB 6; Length 7582;
Best Local Similarity 93.7%; Pred. No. 5.5e-22;
Matches 89; Conservative 6; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCTGCTTATGCGCAAGCTCTTCAGAGAACAAAGAGCCATTACCTCGAGAAGAC 60
|||||
DB 4657 CCTGCTTATGCGCAAGCTCTTCAGAGAACAAAGAGCCATTACCTCGAGAAGAC 4716
|||||

QY 61 CTGGCAACTGATTTTACCACCAAGCCCAAACTCA 95
|||||
DB 4717 CTGGCAACTGATTTTACCACCAAGCCCAAACTCA 4751
|||||

RESULT 12
LOCUS AX027480 7582 bp DNA linear PAT 16-SEP-2000
DEFINITION Sequence 30 from Patent FR2788784.
ACCESSION AX027480
VERSION AX027480.1 GI:10188444
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 7582)
AUTHORS Mallet,F., Voisset,C. and Paranhos,B.G.
JOURNAL Patent: FR 2788784-A 30 28-JUL-2000;
FEATURES
source location/Qualifiers
1..7582
BASE COUNT 2156 a 1876 c 1538 g 1796 t 216 others
ORIGIN

Query Match 97.5%; Score 92.6; DB 6; Length 7582;
Best Local Similarity 93.7%; Pred. No. 5.5e-22;
Matches 89; Conservative 6; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCTGCTTATGCGCAAGCTCTTCAGAGAACAAAGAGCCATTACCTCGAGAAGAC 60
|||||
DB 4657 CCTGCTTATGCGCAAGCTCTTCAGAGAACAAAGAGCCATTACCTCGAGAAGAC 4716
|||||

QY 61 CTGGCAACTGATTTTACCACCAAGCCCAAACTCA 95
|||||
DB 4717 CTGGCAACTGATTTTACCACCAAGCCCAAACTCA 4751
|||||

RESULT 13
LOCUS BD003152 374 bp DNA linear PAT 31-JAN-2002
DEFINITION Secretory proteins and polynucleotides encoding the same.
ACCESSION BD003152
VERSION BD003152.1 GI:18631113
KEYWORDS
SOURCE unidentified.
ORGANISM unidentified.
REFERENCE 1 (bases 1 to 374)
AUTHORS McCoy,J.M., Lavallie,E.R., Racie,L.A., Merberg,D., Treacy,M.,
Evans,C. and Spaulding,V.
TITLE Secretory proteins and polynucleotides encoding the same
JOURNAL Patent: JP 2001501813-A 32 13-FEB-2001;
COMMENT GENETICS INSTITUTE INC, KENNETH JACOBS
OS Unidentified
PN JP 2001501813-A/32
PD 13-FEB-2001

```

PF      23-JUL-1997 JP 1998508939
PR      26-JUL-1996 US 08/686878
PI      JOHN M MCCOY, EDWARD R LEVALLIE, LISA A RACIE, DAVID MERBERG, PI
        MAURICE TREACY,
        PI      CHERYL EVANS, YIKKI SPUNDLING
PC      C12N15/09,A61K38/00,A61K48/00,A61P3/02,A61P7/02,A61P7/08, PC
A61P35/00,
PC      A61P97/00,C07K14/47,C12N5/10,C12P21/02,C12N15/00,C12N5/00, PC
A61K3/02
CC      Strandedness: Double;
CC      Topology: Linear;
FH      key
FT      source
        1..374
        Location/Qualifiers
        1..374
        /organism='Unidentified'
        /db_xref='taxon:32644'
FEATURES
source
        1..374
        Location/Qualifiers
        1..374
        /organism='Unidentified'
BASE COUNT      92 a      100 c      80 g      80 t      22 others
ORIGIN

```

	Query Match	Best Local Similarity	95.2%	Score 90.4	DB 6	Length 374
Matches	91	Conservative	0	Mismatches	3	Indels
						Gaps
						0

LOCUS	180071	180071	374 bp	DNA	linear	PAT 10-JUN-1998
DEFINITION	Sequence 47 from patent US 5708157.					
ACCESSION	180071					
VERSION	180071.1	GI:3208361				
KEYWORDS						
SOURCE	Unknown.					
ORGANISM	Unclassified.					
REFERENCE	1 (bases 1 to 374)					
AUTHORS	Jacobz,K., McCoy,J.M., Lavaille,E.R., Racie,L.A., Merberg,D., Treacy,M., Evans,C. and Spaulding,V.					
TITLE	Secreted proteins and polynucleotides encoding them					
JOURNAL	Patent: US 5708157-A 47 13-JUN-1998;					
FEATURES	Location/Qualifiers					
SOURCE	1..374					

BASE COUNT	92 a	100 c	80 g	80 t	22 others
ORIGIN					
Query Match	95.2%; Score 90.4; DB 6; Length 374;				
Best Local Similarity	96.8%; Pred. No. 2e-21;				
Matches 91; Conservative	0; Mismatches 3; Indels 0; Gaps 0;				
QY	2	CTGCCTATCGCAAGCTCCTTCAGGAGAAACAAGGCCATTACCTGTGAGAGAC	61		
Db	29	CTGCCTATCGCCCAAGCTCTTCAGGAGAACAAAGAAGCAGGCATTACCTGGAAAANAC	88		
QY	62	TGGCAACTGATTTTTACCACCAAGCCCCAACCCTCA	95		
Db	89	TGGCAACTGATTTTTACCACCAAGCCCNAACCTCA	122		
RESULT 15					
LOCUS	AX007978				
DEFINITION	AX007978	2599 bp	DNA	linear	PAT 06-SEP-2000
ACCESSION	AX007978	Sequence 1 from Patent WO9967395.			

VERSION	AX007978.1	GI:995675
KEYWORDS		
SOURCE	human.	
ORGANISM	Homo sapiens	
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.	
AUTHORS	1 (bases 1 to 2599)	
TITLE	Perin,J.P., Rieger,F. and Alliel,P.M.	
JOURNAL	Nucleic sequence and deduced protein sequence family with human endogenous retroviral motifs, and their uses	
FEATURES	Patent: WO 967395-A 1 29-DEC-1999;	
source	INST NAT SANTE RECH MED (FR); PERIN JEAN PIERRE (FR); RIEGER FRANCOIS (FR); ALLIEL PATRICK M (FR)	
	location/Qualifiers	
	1..2599	
	/organism="Homo sapiens"	
	/db_xref="taxon:9606"	
BASE COUNT	744 a	718 c 495 g 642 t
ORIGIN		

	Query Match	Best Local Similarity	Matches	Conservative	Score 87.8%	Pred. No.2.4e-20	DB 6	length 2599	Mismatches 2	Indels 0	Gaps 0
Qy	5	CCTTTGCCCAAGCCTTC	CAGTAAACAACAAGCCATTAC	CCCTGAGAAAGACTGG	64						
Db	10	CTTAATCGCCAGCTCTTC	CAGAGACACAAGAACAGGCCATTAC	CTCGTGGAGACACTGG	69						
Qy	65	CACGTGATTTTACCCACA	AGCCCAAACTCTCA	95							
Db	70	CACGTGATTTTACCCACA	AGCCCAAACTCTCA	100							

Search completed: May 2, 2003, 12:54:41
Job time : 462.862 secs

GenCore version 5.1.5
Copyright (c) 1993 - 2003 CompuGen Ltd

OM nucleic - nucleic search, using sw model

Run on: May 2, 2003, 11:21:20 ; Search time 151.345 Seconds

1413.592 Million cell updates/sec

Title: US-09-719-554-3_COPY_6956_7050

Sequence: 1 cctgccttatcgccaagctc.....accacaagcccaactca 95

Scoring table: IDENTITY_NUC

Searched: 2185239 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters: 4370478

```
Minimum DB seq length: 0
```

Post-1990s

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : N_Geneseq_101002:*

1:	/SID2/gcgdata/genseq/genseqn-emb1/NA1981.DAT *
2:	/SID2/gcgdata/genseq/genseqn-emb1/NA1981.DAT *
3:	/SID2/gcgdata/genseq/genseqn-emb1/NA1982.DAT *
4:	/SID2/gcgdata/genseq/genseqn-emb1/NA1983.DAT *
5:	/SID2/gcgdata/genseq/genseqn-emb1/NA1984.DAT *
6:	/SID2/gcgdata/genseq/genseqn-emb1/NA1985.DAT *
7:	/SID2/gcgdata/genseq/genseqn-emb1/NA1986.DAT *
8:	/SID2/gcgdata/genseq/genseqn-emb1/NA1987.DAT *
9:	/SID2/gcgdata/genseq/genseqn-emb1/NA1988.DAT *
10:	/SID2/gcgdata/genseq/genseqn-emb1/NA1989.DAT *
11:	/SID2/gcgdata/genseq/genseqn-emb1/NA1990.DAT *
12:	/SID2/gcgdata/genseq/genseqn-emb1/NA1991.DAT *
13:	/SID2/gcgdata/genseq/genseqn-emb1/NA1992.DAT *
14:	/SID2/gcgdata/genseq/genseqn-emb1/NA1993.DAT *
15:	/SID2/gcgdata/genseq/genseqn-emb1/NA1994.DAT *
16:	/SID2/gcgdata/genseq/genseqn-emb1/NA1995.DAT *
17:	/SID2/gcgdata/genseq/genseqn-emb1/NA1996.DAT *
18:	/SID2/gcgdata/genseq/genseqn-emb1/NA1997.DAT *
19:	/SID2/gcgdata/genseq/genseqn-emb1/NA1998.DAT *
20:	/SID2/gcgdata/genseq/genseqn-emb1/NA1999.DAT *
21:	/SID2/gcgdata/genseq/genseqn-emb1/NA2000.DAT *
22:	/SID2/gcgdata/genseq/genseqn-emb1/NA2001A.DAT *
23:	/SID2/gcgdata/genseq/genseqn-emb1/NA2001B.DAT *
24:	/SID2/gcgdata/genseq/genseqn-emb1/NA2002.DAT *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	95	100.0	3372	20	AAX25663	Human endogenous r
2	95	100.0	3372	21	AAAS9213	Partial pol gene A
3	95	100.0	3831	23	AA571727	DNA encoding novel
4	95	100.0	5154	23	AA567609	DNA encoding novel
5	95	100.0	7466	23	AA566826	DNA encoding novel
6	95	100.0	8279	23	AA576474	DNA encoding novel
7	95	100.0	8294	23	AA584209	DNA encoding novel
8	95	100.0	10469	21	ABN97929	DNA encoding novel
9	95	100.0	56093	24	ABL61744	Human retroviral s Colon adenocarcinoma

10	94	98.9	2930	24	AAD24195	Human syncytiu
11	94	98.9	2946	21	AAZ77556	Human secreted
12	94	98.9	2946	21	AAZ59468	Human secreted
13	92.6	97.5	7582	20	AAZ52665	Complete human
14	92.6	97.5	7582	21	AAAZ59215	Human endogeno
15	90.4	95.2	374	19	AAAV15135	Human adult tes
16	90.4	95.2	374	19	AAAV09123	5' nucleotide se
17	87.8	92.4	2599	21	ABN97927	Human retrovira
18	85.4	89.9	2006	20	AAAZ5658	Human endogeno
19	85.4	89.9	2006	21	AAAZ59208	Human endogeno
20	81.2	85.5	857	24	ABLI80385	Human ovartian
21	76	80.0	1740	21	ABN97973	Human retrovira
22	74.8	78.7	2575	20	AAAZ5668	Human endogeno
23	74.4	78.3	1218	23	AAZ71171	Human encoding
24	74.4	78.3	1218	23	AAZ84193	DNA encoding
25	74.4	78.3	2585	20	AAZ56567	Human endogeno
26	72.8	76.6	2186	23	AAZ76455	DNA encoding
27	72.8	76.6	4535	19	AAZ76205	DNA encoding
28	71.2	74.9	758	19	AAAZ3222	Multiple scleros
29	71.2	74.9	1158	20	AAAZ9705	Clone LB16 from
30	71.2	74.9	1158	23	AAZ84200	DNA encoding
31	68.2	71.8	1219	23	AAZ73912	DNA encoding
32	63.6	66.9	1216	21	ABN97933	Human retrovira
33	57.4	60.4	2784	21	ABN97930	Human retrovira
34	45.2	47.6	40668	24	ABO88150	Human osteoblast
35	40.4	42.5	747	23	AAO91200	DNA encoding
36	37.2	39.2	1337	23	AAZ66625	DNA encoding
37	37.2	39.2	1983	23	AAZ81202	DNA encoding
38	35.6	37.5	5168	23	AAZ81550	DNA encoding
39	30.8	32.4	414	22	ABAZ5106	Human foetal liv
40	30.8	32.4	414	22	ABAZ59788	Probe #18254 fo
41	30.8	32.4	414	22	AAAZ3623	Human brain exp
42	30.8	32.4	414	22	AAAZ9746	Human bone maro
43	30.8	32.4	414	22	AAZ16884	Probe #16787 fo
44	30.8	32.4	414	22	AAZ15666	Probe #24332 us
45	30.8	32.4	414	24	ABZ53239	Human genome-der

ALIGNMENTS

RESULT 1

ID AAX25663 standard; cDNA to mRNA; 3372 BP.

AC AAX25663;

DT 21-MAY-1999 (first entry)

DE Human endogenous retrovirus W clone cl. P15T.

KW Clone; human endogenous retrovirus; genome; autoimmune disease;

KW disseminated lupus erythematosus; pregnancy; chromosomal marker; ss.

OS Human endogenous retrovirus.

PN W09902696-A1.

PD 21-JAN-1999.

PF 06-JUL-1998;

PR 07-JUL-1997; 97ER-0008815.

PA (INMR) BIO MERIEUX.

PI Beseme F, Blond JL, Bouton O, Mallet F, Mandrand B;

WPI; 1999-120897/10.

PT New nucleic acid sequences from human endogenous retrovirus-W -

PT therapy of autoimmune disease, and abnormal or failed pregnancy
XX
PS Claim 1: Page 64-67; 106pp; French.
XX
CC This sequence represents clone cl.P15f of the human endogenous retrovirus
CC (HERV) W genome. The nucleic acids, their fragments or peptides encoded
CC by them are markers of autoimmune disease (e.g. multiple sclerosis,
CC rheumatoid polyarthritis, disseminated lupus erythematosus, insulin-
CC dependent diabetes and related pathologies) and of abnormal or
CC unsuccessful pregnancy and can be used as chromosomal markers for
CC susceptibility to these conditions, or proximity markers of genes
CC associated with this susceptibility.
XX
SQ Sequence 3372 BP; 1047 A; 835 C; 711 G; 779 T; 0 other;
Query Match 100.0%; Score 95; DB 20; Length 3372;
Best Local Similarity 100.0%; Pred. No. 1.7e-22;
Matches 95; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CCTGCTTATCGCCAGCTCCTTCAGGAGACAAAGAGCGCCATTACCTGGAGAGA 60
DB 1887 CCTGCTTATCGCCAGCTCCTTCAGGAGACAAAGAGCGCCATTACCTGGAGAGA 1946
QY 61 CTGGCACTGATTTTACCCACAAGCCCAACTCA 95
DB 1947 CTGGCACTGATTTTACCCACAAGCCCAACTCA 1981
RESULT 2
ID AAA59213 standard; DNA; 3372 BP.
XX
AC AAA59213;
XX
DT 07-NOV-2000 (first entry)
XX
DE Partial pol gene and U3-R region sequences of HERV-W from human genome.
XX
KW Autoimmune disease; retrovirus; human endogenous retrovirus W; HERV-W;
KW gag gene; pregnancy; multiple sclerosis; T cell proliferation; ss.
XX
OS Homo sapiens.
XX
PN WO200043521-A2.
XX
PD 27-JUL-2000.
XX
PF 21-JAN-2000; 2000WO-FR00144.
XX
PR 21-JAN-1999; 99FR-0000888.
XX
PA (INMR) BIO MERIEUX.
XX
PI Paranhos-Baccala G, Mallet F, Voliset C;
XX
DR WPI: 2000-499229/44.
XX
PT New nucleic acid from human endogenous retrovirus, useful e.g. for
PT diagnosis of autoimmune disease and complications of pregnancy,
PT contains at least part of the gag gene -
XX
XX
PS Disclosure: Page 47-48; 53pp; French.
XX
CC The present sequence represents an endogenous retroviral nucleic acid
CC fragment, which is associated with an autoimmune disease, and is
CC integrated into the human genome. The fragment is originally derived
CC from a novel retrovirus, human endogenous retrovirus W (HERV-W). The
CC HERV-W retrovirus is associated with autoimmune disease, failure of
CC pregnancy or disorders of pregnancy. The nucleic acid fragment, or
CC proteins derived from it, are useful for diagnosis of autoimmune
CC disease (specifically multiple sclerosis) and for monitoring pregnancy.
CC The nucleic acid fragments may also be used for in situ labelling of
CC isolated chromosomes, while the transcription product can be used to

CC study or monitor T cell proliferation in vitro.
XX
SQ Sequence 3372 BP; 1047 A; 835 C; 711 G; 779 T; 0 other;
XX
Query Match 100.0%; Score 95; DB 21; Length 3372;
Best Local Similarity 100.0%; Pred. No. 1.7e-22;
Matches 95; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CCTGCTTATCGCCAGCTCCTTCAGGAGACAAAGAGCGCCATTACCTGGAGAGA 60
DB 1887 CCTGCTTATCGCCAGCTCCTTCAGGAGACAAAGAGCGCCATTACCTGGAGAGA 1946
QY 61 CTGGCACTGATTTTACCCACAAGCCCAACTCA 95
DB 1947 CTGGCACTGATTTTACCCACAAGCCCAACTCA 1981
RESULT 3
ID AAS71727 standard; CDNA; 3831 BP.
XX
AC AAS71727;
XX
DT 13-FEB-2002 (first entry)
XX
DE DNA encoding novel human diagnostic protein #7531.
XX
DE Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX
OS Homo sapiens.
XX
PN WO200175067-A2.
XX
PD 11-OCT-2001.
XX
PF 30-MAR-2001; 2001WO-US08631.
XX
PR 31-MAR-2000; 2000US-0540217.
PR 23-AUG-2000; 2000US-0649167.
XX
PA (HYSE-) HISEQ INC.
XX
PI Drmanac RT, Liu C, Tang YT;
XX
PI P-PSDB; ABC07540.
XX
DR WPI: 2001-639362/73.
XX
PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity -
XX
XX
PS Claim 1: SEQ ID No 7531; 103pp; English.
XX
XX
CC The invention relates to isolated polynucleotide (I) and
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. AAS64197-AAS94364 represent novel human
CC diagnostic coding sequences of the invention.

CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 3831 BP; 1173 A; 953 C; 815 G; 890 T; 0 other;

Query Match 100.0%; Score 95; DB 23; Length 3831;
Best Local Similarity 100.0%; Pred. No. 1.8e-22;
Matches 95; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCGCCTTATCGCCAGCCTCTTCAGAGAAACAAGAGCGCATTTACCTGGAGAGA 60
|||||
DB 2368 CCGCCTTATCGCCAGCCTCTTCAGAGAAACAAGAGCGCATTTACCTGGAGAGA 2427
|||||

QY 61 CTGGCAACTGATTTTACCCACAAGCCCAAACTCA 95
|||||
DB 2428 CTGGCAACTGATTTTACCCACAAGCCCAAACTCA 2462
|||||

RESULT 4

AAS67609
ID AAS67609 standard; cDNA; 5154 BP.

AC AAS67609;

DT 13-FEB-2002 (first entry)

DE DNA encoding novel human diagnostic protein #3413.

XX Human; chromosome mapping; gene mapping; gene therapy; forensic;

KW food supplement; medical imaging; diagnostic; genetic disorder; ss.

XX Homo sapiens.

OS WO200175067-A2.

PN 11-OCT-2001.

PD 30-MAR-2001; 2001WO-US08631.

PF 31-MAR-2000; 2000US-0540217.

PR 23-AUG-2000; 2000US-0649167.

XX (HYSE-) HYSEQ INC.

PA Drmanac RT, Liu C, Tang YT;

PI Drmanac RT, Liu C, Tang YT;

DR WPI; 2001-639362/73.

DR P-PSDB; ABG03422.

XX New isolated polynucleotide and encoded polypeptides, useful in

PT diagnostics, forensics, gene mapping, identification of mutations

PT responsible for genetic disorders or other traits and to assess

PT biodiversity -

PT Claim 1; SEQ ID No 3413; 103bp; English.

PS The invention relates to isolated polynucleotide (I) and

XX polypeptide (II) sequences. (I) is useful as hybridisation probes,

CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome

CC and gene mapping, and in recombinant production of (II). The

CC polynucleotides are also used in diagnostics as expressed sequence tags

CC for identifying expressed genes. (I) is useful in gene therapy techniques

CC to restore normal activity of (II) or to treat disease states involving

CC quantitating a polypeptide in tissue, as molecular weight markers and as

CC a food supplement. (II) and its binding partners are useful in medical

CC imaging of sites expressing (II). (I) and (II) are useful for treating

CC disorders involving aberrant protein expression or biological activity.

CC The polypeptide and polynucleotide sequences have applications in

CC diagnostics, forensics, gene mapping, identification of mutations

CC responsible for genetic disorders or other traits to assess biodiversity

CC and to produce other types of data and products dependent on DNA and

CC amino acid sequences. AAS64197-AAS94564 represent novel human

CC diagnostic coding sequences of the invention.

CC Note: The sequence data for this patent did not appear in the printed

CC specification, but was obtained in electronic format directly from WIPO

CC at ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 5154 BP; 1537 A; 1321 C; 1117 G; 1179 T; 0 other;

Query Match 100.0%; Score 95; DB 23; Length 5154;
Best Local Similarity 100.0%; Pred. No. 2e-22;
Matches 95; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCGCCTTATCGCCAGCCTCTTCAGAGAAACAAGAGCGCATTTACCTGGAGAGA 60
|||||
DB 3691 CCGCCTTATCGCCAGCCTCTTCAGAGAAACAAGAGCGCATTTACCTGGAGAGA 3750
|||||

QY 61 CTGGCAACTGATTTTACCCACAAGCCCAAACTCA 95
|||||
DB 3751 CTGGCAACTGATTTTACCCACAAGCCCAAACTCA 3785
|||||

RESULT 5

AAS68626
ID AAS68626 standard; cDNA; 7466 BP.

AC AAS68626;

DT 13-FEB-2002 (first entry)

DE DNA encoding novel human diagnostic protein #4430.

XX Human; chromosome mapping; gene mapping; gene therapy; forensic;

KW food supplement; medical imaging; diagnostic; genetic disorder; ss.

XX Homo sapiens.

OS WO200175067-A2.

PN 11-OCT-2001.

PD 30-MAR-2001; 2001WO-US08631.

PF 31-MAR-2000; 2000US-0540217.

PR 23-AUG-2000; 2000US-0649167.

XX (HYSE-) HYSEQ INC.

PA Drmanac RT, Liu C, Tang YT;

PI Drmanac RT, Liu C, Tang YT;

DR WPI; 2001-639362/73.

DR P-PSDB; ABG04439.

XX New isolated polynucleotide and encoded polypeptides, useful in

PT diagnostics, forensics, gene mapping, identification of mutations

PT responsible for genetic disorders or other traits and to assess

PT biodiversity -

PT Claim 1; SEQ ID No 4430; 103bp; English.

PS The invention relates to isolated polynucleotide (I) and

XX polypeptide (II) sequences. (I) is useful as hybridisation probes,

CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome

CC and gene mapping, and in recombinant production of (II). The

CC polynucleotides are also used in diagnostics as expressed sequence tags

CC for identifying expressed genes. (I) is useful in gene therapy techniques

CC to restore normal activity of (II) or to treat disease states involving

CC quantitating a polypeptide in tissue, as molecular weight markers and as

CC a food supplement. (II) and its binding partners are useful in medical

CC imaging of sites expressing (II). (I) and (II) are useful for treating

CC disorders involving aberrant protein expression or biological activity.

CC The polypeptide and polynucleotide sequences have applications in

CC diagnostics, forensics, gene mapping, identification of mutations

CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. AAS64197-AAS94564 represent novel human
CC diagnostic coding sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 7466 BP; 2217 A; 1880 C; 1614 G; 1754 T; 1 other;

Query Match 100.0%; Score 95; DB 23; Length 7466;
Best Local Similarity 100.0%; Pred. No. 2.2e-22;
Matches 95; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 CCTGCTTATGCGCAAGCTCTTCAGAGACAAAGAGCCATTACCTGGAGAGA 60
DB 2580 CCTGCTTATGCGCAAGCTCTTCAGAGACAAAGAGCCATTACCTGGAGAGA 2639

OY 61 CTGGCACTGATTTTACCACCAAGCCCAACCTCA 95
DB 2640 CTGGCACTGATTTTACCACCAAGCCCAACCTCA 2674

RESULT 6
AAS76474

ID AAS76474 standard; cDNA; 8279 BP.

AC AAS76474;

DT 13-FEB-2002 (first entry)

DE DNA encoding novel human diagnostic protein #12278.

KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder; ss.

OS Homo sapiens.

PN WO200175067-A2.

PD 11-OCT-2001.

PE 30-MAR-2001; 2001WO-US08631.

PF 31-MAR-2000; 2000US-0540217.

PR 23-AUG-2000; 2000US-0649167.

PA (HYSE-) HYSEQ INC.

PI Drmanac RT, Liu C, Tang YT;

PI WPI; 2001-639362/73.

DR P-PSDB; ABG12287.

PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess

PT biodiversity -
PT Claim 1; SEQ ID No 12278; 103pp; English.

PS The invention relates to isolated polynucleotide (I) and

XX polypeptide (II) sequences. (I) is useful as hybridisation probes,

CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome

CC and gene mapping, and in recombinant production of (II). The

CC polynucleotides are also used in diagnostics as expressed sequence tags

CC for identifying expressed genes. (I) is useful in gene therapy techniques

CC to restore normal activity of (II) or to treat disease states involving

CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.

CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. AAS64197-AAS94564 represent novel human
CC diagnostic coding sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 8279 BP; 2474 A; 1992 C; 1770 G; 2043 T; 0 other;

Query Match 100.0%; Score 95; DB 23; Length 8279;
Best Local Similarity 100.0%; Pred. No. 2.3e-22;
Matches 95; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 CCTGCTTATGCGCAAGCTCTTCAGAGACAAAGAGCCATTACCTGGAGAGA 60
DB 6816 CCTGCTTATGCGCAAGCTCTTCAGAGACAAAGAGCCATTACCTGGAGAGA 6875

OY 61 CTGGCACTGATTTTACCACCAAGCCCAACCTCA 95
DB 6876 CTGGCACTGATTTTACCACCAAGCCCAACCTCA 6910

RESULT 7
AAS84209

ID AAS84209 standard; cDNA; 8294 BP.

AC AAS84209;

DT 13-FEB-2002 (first entry)

DE DNA encoding novel human diagnostic protein #20013.

KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder; ss.

OS Homo sapiens.

PN WO200175067-A2.

PD 11-OCT-2001.

PE 30-MAR-2001; 2001WO-US08631.

PF 31-MAR-2000; 2000US-0540217.

PR 23-AUG-2000; 2000US-0649167.

PA (HYSE-) HYSEQ INC.

PI Drmanac RT, Liu C, Tang YT;

PI WPI; 2001-639362/73.

DR P-PSDB; ABG20022.

PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess

PT biodiversity -
PT Claim 1; SEQ ID No 20013; 103pp; English.

PS The invention relates to isolated polynucleotide (I) and

XX polypeptide (II) sequences. (I) is useful as hybridisation probes,

CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome

CC and gene mapping, and in recombinant production of (II). The

CC polynucleotides are also used in diagnostics as expressed sequence tags

CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical

CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
XX The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations in
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. A564197-A594564 represent novel human
CC diagnostic coding sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at [ftp.wipo.int/pub/published_pct_sequences](http://wipo.int/pub/published_pct_sequences).
SQ Sequence 8294 BP; 2476 A; 1994 C; 1774 G; 2050 T; 0 other;
Query Match 100.0%; Score 95; DB 23; Length 8294;
Best Local Similarity 100.0%; Pred. No. 2.3e-22;
Matches 95; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CCGCCTTATGCGCAGCTCCTTCAGAGAACAAAGACAGCCATTACCTGGAGAAGA 60
DB 6819 CCGCCTTATGCGCAGCTCCTTCAGAGAACAAAGACAGCCATTACCTGGAGAAGA 6878
QY 61 CTGGCAACTGATTTTACCACCAAGCCCAACCTCA 95
DB 6879 CTGGCAACTGATTTTACCACCAAGCCCAACCTCA 6913
RESULT 8
ABN97929
ID ABN97929 standard; DNA: 10499 BP.
XX
AC ABN97929;
XX
DT 01-AUG-2002 (first entry)
XX
DE Human retroviral sequence HERV-7q.
XX
KW Autoimmune disease; HERV-7q; chromosome 7q; immunotherapy;
KM multiple sclerosis; ds.
XX
OS Human retrovirus.
XX
PN WO9967395-A1.
XX
PD 29-DEC-1999.
XX
PF 23-JUN-1999; 99WO-FR01513.
XX
PR 23-JUN-1998; 98FR-0007920.
XX
PA (INRM) INSERM INST NAT SANTE & RECH MEDICALE.
XX
PI Allele PM, Perlin J, Rieger F;
XX
DR WPI; 2000-160587/14.
XX
PT New nucleic acid sequences of human endogenous retrovirus, HERV-7q,
PT used for diagnosis, treatment and prevention of autoimmune and
PT neurological diseases
XX
PS Claim 3; Fig 1; 225pp; French.
XX
CC The present invention relates to new nucleic acid sequences of human
CC endogenous retrovirus, HERV-7q, which is located on chromosome 7q.
CC Regulatory elements associated with HERV-7q may alter expression of other
CC genes (even remote genes) on the same chromosome, inducing immunological
CC and/or neurological changes (which may be pathological or protective/
CC curative). HERV-7q peptides can be used to improve efficiency of the
CC immune response, e.g. in immunotherapy. HERV-7q peptides and their coding
CC sequences can be used in immunogenic or vaccinating compositions, for
CC protection against autoimmune diseases, particularly multiple sclerosis.
CC The peptides may also be used (by sequence comparison) to detect/identify
CC endogenous retroviruses that are abnormally expressed in cancer.

CC neuropathologies or other autoimmune diseases. The present sequence was
CC used to illustrate the invention.
XX
SQ Sequence 10499 BP; 3048 A; 2676 C; 2280 G; 2495 T; 0 other;
Query Match 100.0%; Score 95; DB 21; Length 10499;
Best Local Similarity 100.0%; Pred. No. 2.5e-22;
Matches 95; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CCGCCTTATGCGCAGCTCCTTCAGAGAACAAAGACAGCCATTACCTGGAGAAGA 60
DB 6956 CCGCCTTATGCGCAGCTCCTTCAGAGAACAAAGACAGCCATTACCTGGAGAAGA 7015
QY 61 CTGGCAACTGATTTTACCACCAAGCCCAACCTCA 95
DB 7016 CTGGCAACTGATTTTACCACCAAGCCCAACCTCA 7050
RESULT 9
ABL61744
ID ABL61744 standard; DNA: 56093 BP.
XX
AC ABL61744;
XX
DT 15-MAY-2002 (first entry)
XX
DE Colon adenocarcinoma related gene sequence SEQ ID NO:81.
XX
KW Human; cancer; colon; breast; ovary; oesophagus; kidney; thyroid;
KW stomach; lung; prostate; pancreas; carcinoma; antitumour; cancerous;
KW cytostatic; gene therapy; antineoplastic; Wilms' tumour; adenocarcinoma;
KW gene; ds.
XX
OS Homo sapiens.
XX
PN WO200194629-A2.
XX
PD 13-DEC-2001.
XX
PF 30-MAY-2001; 2001WO-US10838.
XX
PR 05-JUN-2000; 2000US-209473P.
PR 05-JUN-2000; 2000US-209531P.
PR 18-SEP-2000; 2000US-231133P.
PR 18-SEP-2000; 2000US-233617P.
PR 20-SEP-2000; 2000US-234009P.
PR 20-SEP-2000; 2000US-234034P.
PR 20-SEP-2000; 2000US-234052P.
PR 22-SEP-2000; 2000US-234509P.
PR 22-SEP-2000; 2000US-234567P.
PR 23-SEP-2000; 2000US-234923P.
PR 23-SEP-2000; 2000US-234924P.
PR 25-SEP-2000; 2000US-235077P.
PR 25-SEP-2000; 2000US-235082P.
PR 25-SEP-2000; 2000US-235134P.
PR 25-SEP-2000; 2000US-235280P.
PR 26-SEP-2000; 2000US-235637P.
PR 26-SEP-2000; 2000US-235638P.
PR 27-SEP-2000; 2000US-235711P.
PR 27-SEP-2000; 2000US-235720P.
PR 27-SEP-2000; 2000US-235840P.
PR 27-SEP-2000; 2000US-235863P.
PR 28-SEP-2000; 2000US-236028P.
PR 28-SEP-2000; 2000US-236032P.
PR 28-SEP-2000; 2000US-236033P.
PR 28-SEP-2000; 2000US-236034P.
PR 28-SEP-2000; 2000US-236109P.
PR 28-SEP-2000; 2000US-236111P.
PR 28-SEP-2000; 2000US-236842P.
PR 29-SEP-2000; 2000US-236891P.
PR 02-OCT-2000; 2000US-237172P.
PR 02-OCT-2000; 2000US-237173P.
PR 02-OCT-2000; 2000US-237278P.

PR 02-OCT-2000; 2000US-237294P.
PR 02-OCT-2000; 2000US-237295P.
PR 02-OCT-2000; 2000US-237316P.
PR 03-OCT-2000; 2000US-237425P.
PR 03-OCT-2000; 2000US-237598P.
PR 03-OCT-2000; 2000US-237598P.
PR 03-OCT-2000; 2000US-237604P.
PR 03-OCT-2000; 2000US-237606P.
PR 03-OCT-2000; 2000US-237608P.
PR 01-NOV-2000; 2000US-244867P.
PR 01-NOV-2000; 2000US-245084P.
PA (AVALON) AVALON PHARM.
PI Young PE, Augustus M, Carter KC, Ebner R, Endress G, Horrigan S;
PI Soppet DR, Weaver Z;
XX WPI; 2002-188264/24.
XX
XX
XX Screening for anti-neoplastic agent involves exposing cells to a
PT chemical agent to be tested for anti-neoplastic activity, and
PT determining a change in expression of a gene of a signature gene set -
XX
XX Claim 1; SEQ ID 81; 44pp; English.
XX
XX The present invention describes a method (M1) for screening for an
CC anti-neoplastic agent. The method involves exposing cells to a chemical
CC agent to be tested for anti-neoplastic activity, determining a change in
CC expression of at least one gene (I) of a signature gene set, where (I)
CC comprises a sequence (S) selected from 8447 sequences (given in ABL6164
CC to ABL70110), or is at least 95% identical to (S), where a change in
CC expression is indicative of anti-neoplastic activity. (I) has cytostatic
CC activity and can be used in gene therapy. M1 can be used for screening
CC an anti-neoplastic agent, and can be used for producing a product which
CC is the data collected with respect to the anti-neoplastic agent as a
CC result of M1, and the data is sufficient to convey the chemical
CC structure and/or properties of the agent. M1 can be used in the
CC treatment of cancer such as colon, breast, stomach, lung, thyroid,
CC oesophageal, ovarian, kidney, prostate or pancreatic cancer,
CC adenocarcinoma, carcinoma, clear cell cancer, infiltrating ductal cancer,
CC infiltrating lobular cancer, squamous cell carcinoma, neuroendocrine
CC carcinoma, papillary carcinoma and Wilms' tumour.
XX
XX Sequence 56093 BP; 16164 A; 12346 C; 10702 G; 16881 T; 0 other;
SQ
Query Match 100.0%; Score 95; DB 24; Length 56093;
Best Local Similarity 100.0%; Pred. No. 4.1e-22;
Matches 95; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CCTGCTTATCGCAAGCTCTTCAGAGAACAAAGACAGGCCATTACCTGGAGAGA 60
DB 34956 CTTGCTTATCGCAAGCTCTTCAGAGAACAAAGACAGGCCATTACCTGGAGAGA 35015
QY 61 CTGGCAACTGATTATTACCCACAAGCCCAAACTCA 95
DB 35016 CTGGCAACTGATTATTACCCACAAGCCCAAACTCA 35050
RESULT 10
AAD24195
ID AAD24195 standard; cDNA; 2930 BP.
XX
XX AAD24195;
XX
XX 07-MAY-2002 (first entry)
XX
XX Human syncytin cDNA.
XX
XX Human; syncytin; pre-eclampsia; gestational trophoblast disorder;
KW choriocarcinoma; hydatiform mole; placental site tumour; abortion;
KW envelope gene; human endogenous defective retrovirus; HERV-W; ss.
XX
XX Homo sapiens.
XX

FH Key Location/Qualifiers
FT CDS 930..2546
FT /*tag= a
FT /product= "Syncytin"
PN WO200204678-A2.
XX
XX 17-JAN-2002.
PD
XX
XX 09-JUL-2001; 2001WO-US21719.
PF
XX
XX 07-JUL-2000; 2000US-216657P.
XX
XX (GEMV) GENETICS INST INC.
PA
XX
XX Keith JC, McCoy JM, M1 S;
PI
XX
XX WPI; 2002-171727/22.
DR
XX
XX P-PSDB; AAE14540.
DR
XX
XX
XX Identifying a compound for treating a subject with or at risk of
PT developing pre-eclampsia, comprises determining whether the expression
PT or activity of syncytin in the cell is modulated in the presence of a
PT test compound -
XX
XX
XX Disclosure; Page 39-42; 43pp; English.
XX
XX The invention relates to identifying compounds which are modulators
CC of syncytin expression. The syncytin modulators are useful in diagnosis
CC and treatment of pre-eclampsia and gestational trophoblast disorders (e.g.
CC choriocarcinoma, hydatiform mole, placental site tumour and missed/
CC incomplete abortion). Syncytin is a human gene derived from the
CC envelope gene of human endogenous defective retrovirus, HERV-W. The
CC present invention is based partly on the discovery that syncytin
CC expression is dramatically reduced in pre-eclampsia, and is also
CC mis-localised to the apical syncytiotrophoblast membrane. The present
CC sequence is human syncytin cDNA.
XX
XX Sequence 2930 BP; 842 A; 800 C; 571 G; 717 T; 0 other;
SQ
Query Match 98.9%; Score 94; DB 24; Length 2930;
Best Local Similarity 100.0%; Pred. No. 3.7e-22;
Matches 94; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2 CTGCTTATCGCAAGCTCTTCAGAGAACAAAGACAGGCCATTACCTGGAGAGAC 61
DB 8 CTGCTTATCGCAAGCTCTTCAGAGAACAAAGACAGGCCATTACCTGGAGAGAC 67
QY 62 TGGCAACTGATTATTACCCACAAGCCCAAACTCA 95
DB 68 TGGCAACTGATTATTACCCACAAGCCCAAACTCA 101
RESULT 11
AAx77526
ID AAx77526 standard; cDNA; 2946 BP.
XX
XX AAx77526;
XX
XX 10-AUG-1999 (first entry)
XX
XX Human secreted protein AJ172_2 cDNA.
XX
XX Secreted protein; testes; brain; blood; placenta; human; murine; thymus;
KW bone marrow; treatment; prevention; nutrition; cytokine; immune; vaccine;
KW cell proliferation; cell differentiation; suppressor; tumour inhibition;
KW haematopoiesis regulator; activin; inhibin; chemotactic; chemokine;
KW haemostatic; thrombolytic; receptor; ligand; anti-inflammatory; tumour;
KW cadherin; tumour invasion suppressor; gene therapy; tissue growth; ss.
XX
XX Homo sapiens.
XX
XX WO9926972-A1.
XX
XX

[illegible]

XX	17-MAY-1999;	99WO-US10915.
XX	18-MAY-1998;	98US-0080478.
PR	20-OCT-1998;	98US-0175928.
XX	(GENY) GENETICS INST INC.	
PA	Jacobs K, McCoy JM, Lavallic ER, Collins-Racie LA, Evans C;	
PI	Merberg D, Mi S, Treacy M;	
PI	WPI: 2000-116311/10.	
DR	P-FSDB; AAY67313.	
XX	New polynucleotides encoding secreted cDNA libraries, used to develop	
PT	products for the diagnosis and treatment of neoplastic disease	
XX	Claim 14; Page 107-108; 149pp; English.	
PS	This is the human secreted protein A172.2 nucleotide sequence, obtained	
XX	from a human adult testes cDNA library. The invention relates to secreted	
CC	human and murine proteins. The polynucleotides and proteins are predicted	
CC	to have biological activities which would make them suitable for	
CC	treating, preventing or ameliorating medical conditions in humans and	
CC	animals. Detection of the levels of the proteins can be used for the	
CC	diagnosis of e.g. pre-eclampsia, placental pathology or cancer. Agents	
CC	which modulate the expression or function of the proteins may be used for	
CC	treating a neoplastic disease and inhibiting metastasis. Other suggested	
CC	activities include nutritional activity (e.g. in feeds), cytokine and	
CC	cell proliferation/differentiation activity, immune stimulating	
CC	(e.g. as vaccines) or suppressing activity, haematopoiesis regulating	
CC	activity, tissue growth activity, activin/inhibin activity,	
CC	chemotactic/chemokinetic activity, haemostatic and thrombolytic activity,	
CC	receptor/ligand activity, anti-inflammatory activity, cadherin/tumour	
CC	invasion suppressor activity, and tumour inhibition activity. The	
CC	polynucleotide sequences are also stated to be useful for gene therapy.	
XX	Sequence 2946 BP; 858 A; 801 C; 570 G; 717 T; 0 other;	
XX	Query Match	98.9%; Score 94; DB 21; Length 2946;
XX	Best Local Similarity	100.0%; Pred. No. 3.7e-22;
XX	Matches	94; Conservative 0; Mismatches 0; Indels 0; Gaps 0
QY	2 CTGCTTATCGCAAGCTCTTCAGAGACAACAAGACAGCCATTACCTGGAGAAGAC	61
DB	6 CTGCCTTATCGCAGCAAGCTCTTCAGAGACAACAAGACAGCCATTACCTGGAGAAGAC	65
QY	62 TGGCAACTGATTTTACCCACAAGCCCAAACTCA	95
DB	66 TGGCAACTGATTTTACCCACAAGCCCAAACTCA	99
XX	RESULT 13	
XX	AAAX25665	
ID	AAAX25665 standard; cDNA to mRNA; 7582 BP.	
XX	AAAX25665;	
XX	21-MAY-1999 (first entry)	
DE	Complete human endogenous retrovirus W genome.	
XX	Clone; human endogenous retrovirus; genome; autoimmune disease;	
KW	multiple sclerosis; rheumatoid polyarthritis; insulin-dependent diabetes;	
XX	disseminated lupus erythematosus; pregnancy; chromosomal marker; ss.	
XX	Human endogenous retrovirus.	
OS	WO9902696-A1.	
PN	21-JAN-1999.	
XD	06-JUL-1998; 98WO-FR01442.	
XX		

XX 07-JUL-1997: 97FR-0008815.
 PR (INMR) BIO MERIEUX.
 PA
 XX Beseme F, Blond JL, Bouton O, Mallet F, Mandrand B;
 PI WPI; 1999-120897/10.
 DR
 XX New nucleic acid sequences from human endogenous retrovirus-W -
 PT expressed exclusively in placenta and useful in diagnosis and
 PT therapy of autoimmune disease, and abnormal or failed pregnancy
 XX
 PS Claim 1; Page 71-74; 106pp; French.
 XX
 CC This sequence represents the complete sequence of the human endogenous
 CC retrovirus (HERV) W genome. The nucleic acids, their fragments or
 CC peptides encoded by them are markers of autoimmune disease (e.g. multiple
 CC sclerosis, rheumatoid polyarthritis, disseminated lupus erythematosus,
 CC insulin-dependent diabetes and related pathologies) and of abnormal or
 CC unsuccessful pregnancy and can be used as chromosomal markers for
 CC susceptibility to these conditions, or proximity markers of genes
 CC associated with this susceptibility.
 XX
 SQ Sequence 7582 BP; 2156 A; 1877 C; 1537 G; 1796 T; 2 U; 214 other;
 Query Match 97.5%; Score 92.6; DB 20; Length 7582;
 Best Local Similarity 93.7%; Pred. No. 1.5e-21;
 Matches 89; Conservative 6; Mismatches 0; Indels 0; Gaps 0;
 YY
 CC 1 CCTGCTTATGCGCAAGCTCTCTTACAGAGAACAAGACGACCATTAACCTGAGAGA 60
 DB 4657 CCTGCTTATGCGCAAGCTCTCTTACAGAGAACAAGACGACCATTAACCTGAGAGA 4716
 YY 61 CTGGCAACTGATTTTACCACCAAGCCCAAACTCA 95
 DB 4717 CTGGCAACTGATTTTACCACCAAGCCCAAACTCA 4751
 RESULT 14
 ID AAV59215 standard; DNA: 7582 BP.
 XX
 AC AAV59215;
 XX
 DT 07-NOV-2000 (first entry)
 DE Human endogenous retrovirus W (HERV-W) sequence.
 XX
 KW Autoimmune disease; retrovirus; human endogenous retrovirus W; HERV-W;
 KW gag gene; pregnancy; multiple sclerosis; T cell proliferation; ss.
 XX
 OS Human endogenous retrovirus.
 XX
 FH Key Location/Qualifiers
 FT LTR 1..120
 FT /*tag- a
 FT /note- "R of 5' LTR"
 FT LTR 121..575
 FT /*tag- b
 FT /note- "U5 of 5' LTR"
 FT primer_bind 579..596
 FT /*tag- c
 FT CDS 581..7194
 FT /*tag- d
 FT /note- "ORF1 env538"
 FT CDS 7039..7194
 FT /*tag- e
 FT CDS 7112..7255
 FT /note- "ORF2 52 AA"
 FT /*tag- f
 FT /note- "ORF3 48 AA"
 FT misc_feature 7244..7254

FT /*tag- g
 FT /note- "polypurine tract"
 FT LTR 7256..7582
 FT /*tag- h
 FT /note- "U3-R of 3' LTR"
 FT polyA_signal 7563..7569
 FT /*tag- i
 PN W0200043521-A2.
 XX
 PD 27-JUL-2000.
 XX
 PF 21-JAN-2000; 2000WO-FR00144.
 XX
 PR 21-JAN-1999; 99FR-0000888.
 XX
 PA (INMR) BIO MERIEUX.
 PI Paranhos-Baccala G, Mallet F, Voliset C;
 DR WPI; 2000-499229/44.
 XX
 XX New nucleic acid from human endogenous retrovirus, useful e.g. for
 PT diagnosis of autoimmune disease and complications of pregnancy,
 PT contains at least part of the gag gene
 XX
 PS Disclosure; Page 49-52; 53pp; French.
 XX
 CC The present sequence represents an endogenous retrovirus, which is
 CC associated with an autoimmune disease, and is integrated into the human
 CC genome. The retrovirus is human endogenous retrovirus W (HERV-W). The
 CC HERV-W retrovirus is associated with autoimmune disease, failure of
 CC pregnancy or disorders of pregnancy. HERV-W nucleic acid fragments, or
 CC proteins derived from it, are useful for diagnosis of autoimmune
 CC disease (specifically multiple sclerosis) and for monitoring pregnancy.
 CC The nucleic acid fragments may also be used for in situ labelling of
 CC isolated chromosomes, while the transcription product can be used to
 CC study or monitor T cell proliferation in vitro.
 XX
 SQ Sequence 7582 BP; 2156 A; 1876 C; 1538 G; 1796 T; 216 other;
 Query Match 97.5%; Score 92.6; DB 21; Length 7582;
 Best Local Similarity 93.7%; Pred. No. 1.5e-21;
 Matches 89; Conservative 6; Mismatches 0; Indels 0; Gaps 0;
 YY
 CC 1 CCTGCTTATGCGCAAGCTCTCTTACAGAGAACAAGACGACCATTAACCTGAGAGA 60
 DB 4657 CCTGCTTATGCGCAAGCTCTCTTACAGAGAACAAGACGACCATTAACCTGAGAGA 4716
 YY 61 CTGGCAACTGATTTTACCACCAAGCCCAAACTCA 95
 DB 4717 CTGGCAACTGATTTTACCACCAAGCCCAAACTCA 4751
 RESULT 15
 ID AAV15135 standard; CDNA: 374 BP.
 XX
 AC AAV15135;
 XX
 DT 29-JUN-1998 (first entry)
 DE Human adult testes cDNA clone A172_2 5' cDNA portion.
 XX
 KW Human; adult; testes; A172_2; secreted protein; cytokine; nutrition;
 KW cell proliferation; differentiation; tumour; immune; ds.
 XX
 OS Homo sapiens.
 XX
 PN W09804696-A1.
 XX
 PD 05-FEB-1998.
 XX

```

PF 23-JUL-1997; 97WO-US12910.
XX
PR 27-SEP-1996; 96US-0721489.
PR 26-JUL-1996; 96US-0686878.
XX
PA (GEMT ) GENETICS INST INC.
XX
PI Jacobs K, Lavallie ER, McCoy JM, Merberg D, Racie LA;
PI Spaulding V, Treacy M;
XX
DR WPI; 1998-130687/12.
XX
PT Polynucleotides and secreted proteins having cytokine and cell
PT proliferation/differentiation activity - isolated from a human adult
PT testes cDNA library
XX
PS Claim 13; Page 33; 48pp; English.
XX
CC The present sequence represents the 5' cDNA portion of the
CC clone AJ172.2 from a human adult testes cDNA library. The nucleotide
CC sequences and protein from clone AJ172.2, have uses in nutrition,
CC cytokine and cell proliferation/differentiation activity, immune
CC stimulating or suppressing activity, hematopoiesis regulating activity,
CC chemotactic/chemokinetic activity, haemostatic or thrombotic activity,
CC receptor/ligand activity, anti-inflammatory activity, tumour inhibition
CC activity or other activities. The products can also be used for
CC production of antibodies and for detection and diagnosis.
XX
SQ Sequence 374 BP; 92 A; 100 C; 80 G; 80 T; 22 other;

```

Query Match 95.2%; Score 90.4; DB 19; Length 374;
Best Local Similarity 96.8%; Pred. No. 3.4e-21;
Matches 91; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```

QY 2 CTGCTTATGCGCAGCTCTTCAGGAGAACAAAGACAGGCCATTACCTGGAGAGAC 61
   |||||||
DB 29 CTGCTTATGCGCAGCTCTTCAGGAGAACAAAGACAGGCCATTACCTGGAGAGAC 88
   |||||||
QY 62 TGGCACTGATTTTACCCACAGCCCAACCTCA 95
   |||||||
DB 89 TGGCACTGATTTTACCCACAGCCCAACCTCA 122
   |||||||

```

Search completed: May 2, 2003, 11:30:20
Job time : 174.345 secs

GenCore version 5.1.5
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 2, 2003, 11:21:20 ; Search time 1094.14 Seconds

(without alignments)
1406.195 Million cell updates/sec

Title: US-09-719-554-3_COPY_6956_7050

Perfect score: 95
Sequence: 1 cctgcctatgcgcagctc.....accacaagcccaactcctca 95

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estnu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hlc:*
9: gb_estl:*
10: gb_estl2:*
11: gb_hlc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pln:*
21: em_gss_yrt:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_trod:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	95	100.0	372	12	BF992537 IL5-GN017
2	95	100.0	379	12	BF993049 IL5-GN017
3	95	100.0	423	12	BF992535 IL5-GN017
4	95	100.0	531	12	BG012022 QV3-GN033
5	81.2	85.5	337	9	AA362784 EST72629
6	77.6	81.7	738	17	HSA408776 Homo sapi

7	70	73.7	723	13	BT72113	BT72113 603059196
8	67.8	71.4	452	12	BF761543	BF761543 CMO-CS004
9	59	62.1	321	9	AA179079	AA179079 zp1le01.r
10	55.6	58.5	508	9	AA176759	AA176759 zp1le01.s
11	42.4	44.6	660	17	AQ469362	AQ469362 CITBI-E1-
12	37.2	39.2	385	17	AQ109046	AQ109046 CIT-HSP-2
13	37.2	39.2	439	17	AQ196669	AQ196669 CIT-HSP-2
14	37.2	39.2	604	14	BM713850	BM713850 UT-E-EJ0-
15	35.2	37.1	489	17	B93732	B93732 CIT-HSP-216
16	34.2	36.0	188	13	BG960113	BG960113 PM3-CT064
17	33.4	35.2	540	17	AQ610340	AQ610340 HS-5093-A
18	33.2	34.9	535	17	AQ153245	AQ153245 HS-2230-B
19	32.8	34.5	674	17	AZ870530	AZ870530 2M0183J06
20	32.6	34.3	165	13	BG960120	BG960120 PM3-CT064
21	30	31.6	519	17	AQ753746	AQ753746 HS-2137-B
22	29.4	30.9	383	12	BG406840	BG406840 d4ac20G05.
23	29.4	30.9	404	12	BF557280	BF557280 UT-R-C0-g
24	29.2	30.7	602	17	AQ234153	AQ234153 HS-2057_A
25	29.2	30.7	703	17	AZ853642	AZ853642 2M0156M23
26	29	30.5	676	17	AQ053323	AQ053323 CIT-HSP-2
27	29	30.5	849	17	AZ898632	AZ898632 RRC1-24-1
28	28.8	30.3	547	17	AZ128099	AZ128099 OSUNBD009
29	28.8	30.3	626	17	BB636193	BB636193 BB636193
30	28.8	30.3	755	13	BJ062473	BJ062473 RC2-HT077
31	28.6	30.1	344	12	BE717297	BE717297 RC2-HT077
32	28.6	30.1	522	10	BE341820	BE341820 EST94654
33	28.6	30.1	597	12	BG593585	BG593585 EST492263
34	28.6	30.1	657	13	BM407602	BM407602 EST581917
35	28.6	30.1	703	12	BG590087	BG590087 EST497929
36	28.4	29.9	289	13	B1120199	B1120199 F011P78Y
37	28.4	29.9	324	17	AZ628023	AZ628023 1M0476A14
38	28.4	29.9	298	13	B1120140	B1120140 F011P07Y
39	28.4	29.9	353	13	B1119806	B1119806 F005P50Y
40	28.4	29.9	379	10	AM276118	AM276118 xq45h08.x
41	28.4	29.9	381	13	B1138741	B1138741 F114P91Y
42	28.4	29.9	473	13	B1127583	B1127583 G062P66Y
43	28.4	29.9	474	13	B1138181	B1138181 F100P47Y
44	28.4	29.9	490	9	AT161614	AT161614 A004P03U
45	28.4	29.9	517	13	BT131573	BT131573 G122P74Y

ALIGNMENTS

RESULT 1
LOCUS BF992537 372 bp mRNA linear EST 23-JAN-2001
DEFINITION IL5-GN0178-301000-203-f06 GN0178 Homo sapiens CDNA, mRNA sequence.
ACCESSION BF992537
VERSION BF992537.1 GI:12398860
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens

REFERENCE
AUTHORS
Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R., Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F., Goldman,G.H., Carvalho,A.F., Matsukuma,A., Bala,G.S., Simpson,D.H., Brunstein,A., deoliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and Simpson,A.J.J.
Shotgun sequencing of the human transcriptome with ORF expressed sequence tags
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
CONTACT: Simpson A.J.J.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001

TITLE
JOURNAL
MEDLINE
COMMENT

Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?l=IL5&t2=IL5-GN0178-301000-203-106&t3=2000-10-30&t4=1)
Seq primer: puc 18 forward
High quality sequence stop: 371.

FEATURES

Location/Qualifiers
1..372

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="GN0178"
/dev_stage="Adult"

/note="Organ: placenta_normal; Vector: puc18; Site_1: SmaI
; Site_2: SmaI; A mini-library was made by cloning
products derived from ORESTES PCR (U.S. Letters Patent
application No. 196,716 - Ludwig Institute for Cancer
Research) profiles into the pUC 18 vector. Reverse
transcription of tissue mRNA and cDNA amplification were
performed under low stringency conditions."

BASE COUNT

116 a 93 c 76 g 86 t 1 others

ORIGIN

Query Match 100.0%; Score 95; DB 12; Length 372;
Best Local Similarity 100.0%; Pred. No. 1.1e-20;
Matches 95; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY

1 CCTGCTTATCGCCAGCTCTTTCAGAGAAAGAGCCGATTACCTGGAGAGA 60

Db

179 CTTGCTTATCGCCAGCTCTTTCAGAGAAAGAGCCGATTACCTGGAGAGA 238

QY

61 CTGGCAACTGATTTTACCACCAAGCCCAAACTCA 95

Db

239 CTGGCAACTGATTTTACCACCAAGCCCAAACTCA 273

RESULT 2

BF993049 379 bp mRNA linear EST 23-JAN-2001
LOCUS IL5-GN0178-301000-204-d04 GN0178 Homo sapiens cDNA, mRNA sequence.
DEFINITION BF993049

ACCESSION

BF993049

VERSION

BF993049.1 GI:12399372

KEYWORDS

EST.

SOURCE

human.

ORGANISM

Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

REFERENCE

Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 379)

AUTHORS

Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
Nagal,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
Goldman,G.H., Carvalho,A.F., Matsukuma,A., Bala,G.S., Simpson,D.H.,
Brunstein,A., deoliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
Simpson,A.J.

TITLE

Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags

JOURNAL

Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)

MEDLINE

20202663

COMMENT

Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil

JOURNAL

Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)

MEDLINE

20202663

COMMENT

Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil

JOURNAL

Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)

MEDLINE

20202663

COMMENT

Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil

JOURNAL

Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)

MEDLINE

20202663

COMMENT

Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil

JOURNAL

Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)

MEDLINE

20202663

source
1..379
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="GN0178"
/dev_stage="Adult"

BASE COUNT

118 a 94 c 77 g 89 t 1 others

ORIGIN

Query Match 100.0%; Score 95; DB 12; Length 379;
Best Local Similarity 100.0%; Pred. No. 1.1e-20;
Matches 95; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY

1 CTTGCTTATCGCCAGCTCTTTCAGAGAAAGAGCCGATTACCTGGAGAGA 60

Db

179 CTTGCTTATCGCCAGCTCTTTCAGAGAAAGAGCCGATTACCTGGAGAGA 238

QY

61 CTGGCAACTGATTTTACCACCAAGCCCAAACTCA 95

Db

239 CTGGCAACTGATTTTACCACCAAGCCCAAACTCA 273

RESULT 3

BF992535 423 bp mRNA linear EST 23-JAN-2001
LOCUS IL5-GN0178-301000-203-e12 GN0178 Homo sapiens cDNA, mRNA sequence.
DEFINITION BF992535

ACCESSION

BF992535

VERSION

BF992535.1 GI:12398858

KEYWORDS

EST.

SOURCE

human.

ORGANISM

Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

REFERENCE

Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 423)

AUTHORS

Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
Nagal,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
Goldman,G.H., Carvalho,A.F., Matsukuma,A., Bala,G.S., Simpson,D.H.,
Brunstein,A., deoliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
Simpson,A.J.

TITLE

Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags

JOURNAL

Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)

MEDLINE

20202663

COMMENT

Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil

JOURNAL

Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)

MEDLINE

20202663

COMMENT

Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil

JOURNAL

Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)

MEDLINE

20202663

COMMENT

Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil

JOURNAL

Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)

MEDLINE

20202663

COMMENT

Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil

JOURNAL

Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)

MEDLINE

20202663

COMMENT

Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil

JOURNAL

Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)

MEDLINE

20202663

COMMENT

Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil

JOURNAL

Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)

MEDLINE

20202663

COMMENT

Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil

FEATURES

Location/Qualifiers
1..423

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="GN0178"
/dev_stage="Adult"

/note="Organ: placenta_normal; Vector: puc18; Site_1: SmaI
; Site_2: SmaI; A mini-library was made by cloning
products derived from ORESTES PCR (U.S. Letters Patent

application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

BASE COUNT

129 a 108 c 88 g 98 t

Query Match 100.0%; Score 95; DB 12; Length 423;
Best Local Similarity 100.0%; Pred. No. 1.1e-20;
Matches 95; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCTGCTTATGCCAAGCTCTTTCAGAGAACAAAGAGCCATTACCTGTGAGAGA 60
DB 179 CCTGCTTATGCCAAGCTCTTTCAGAGAACAAAGAGCCATTACCTGTGAGAGA 238
QY 61 CTGGCACTGATTTTACCCACAGCCCAACCTCA 95
DB 239 CTGGCACTGATTTTACCCACAGCCCAACCTCA 273

RESULT 4

LOCUS BG012022 531 bp mRNA linear EST 24-JAN-2001
DEFINITION QV3-GN0335-131200-541-d12 GN0335 Homo sapiens CDNA, mRNA sequence.
ACCESSION BG012022
VERSION BG012022.1 GI:12460813
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens

REFERENCE

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 531)
Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R., Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F., Goldman, G.H., Carvalho, A., Matsukuma, A., Bala, G.S., Simpson, D.H., Brunstein, A., deoliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare, M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and Simpson, A.J.

Shotgun sequencing of the human transcriptome with ORF expressed

sequence tags

Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)

COMMENT

20202663
Laboratory of Simpson A.J.G.
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br

This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?l=QV3&t2=QV3-GN0335-131200-541-d12&t3=2000-12-13&t4=1)
Seq primer: puc 18 forward
High quality sequence start: 18
High quality sequence stop: 530.

FEATURES

Source

1..531

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="GN0335"

/dev_stage="Adult"

/note="Organ: Placenta,normal; Vector: puc18; Site_1: Sma1; Site_2: Sma1; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent Application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

BASE COUNT

110 a 117 c 139 g 165 t

Query Match 100.0%; Score 95; DB 12; Length 531;
Best Local Similarity 100.0%; Pred. No. 1.2e-20;
Matches 95; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCTGCTTATGCCAAGCTCTTTCAGAGAACAAAGAGCCATTACCTGTGAGAGA 60
DB 405 CCTGCTTATGCCAAGCTCTTTCAGAGAACAAAGAGCCATTACCTGTGAGAGA 346
QY 61 CTGGCACTGATTTTACCCACAGCCCAACCTCA 95
DB 345 CTGGCACTGATTTTACCCACAGCCCAACCTCA 311

RESULT 5

LOCUS AA362784/c 337 bp mRNA linear EST 21-APR-1997
DEFINITION EST72629 Ovary II Homo sapiens cDNA 5' end, mRNA sequence.
ACCESSION AA362784
VERSION AA362784.1 GI:2015175
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens

REFERENCE

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 337)
Adams, M.D., Kerlavage, A.R., Fleischmann, R.D., Fulmer, R.A., Bult, C.J., Lee, N.H., Kinkness, E.F., Weissbrock, K.G., Gocayne, J.D., White, O., Sutton, G., Blake, J.A., Brandon, R.C., Man, W.-C., Clayton, R.A., Cline, T.R., Cotton, M.D., Earle-Hughes, J., Fine, L.D., Fitzgerald, L.M., Fitzhugh, W.M., Fritchman, J.L., Geoghegan, N.S., Glodek, A., Gnehm, C.L., Hanna, M.C., Hedblom, E., Hinkle, P.S. Jr., Kelley, J.M., Kelley, J.C., Liu, L.-I., Marmaro, S.M., Merrick, J.M., Moreno-Palauques, R.F., McDonald, L.A., Nguyen, D.T., Pelligrino, S.M., Phillips, C.A., Ryder, S.E., Scott, J.L., Saudak, D.M., Shirley, R., Small, K.V., Spriggs, T.A., Utterback, T.R., Weidman, J.F., Li, Y., Bednarek, D.P., Cao, L., Cepeda, M.A., Coleman, T.A., Collins, E.J., Diabe, D., Feng, D.-F., Ferris, A., Fischer, C., Hastings, G.A., He, W.W., Hu, J.S., Greene, J.M., Gruber, J., Hudson, P., Kim, A.K., Kozak, D.L., Kunsch, C., Hungjun, J., Li, H., Weissner, P.S., Olsen, H., Raymond, L., Wei, Y.F., Wing, J., Xu, C., Yu, G.L., Ruben, S.M., Dillion, P.J., Fannon, M.R., Rosen, C.A., Haseltine, W.A., Fields, C., Fraser, C.M. and Venter, J.C.

TITLE

Initial assessment of human gene diversity and expression patterns based upon 83 million nucleotides of cDNA sequence

Nature 377 (6547 Suppl), 3-174 (1995)

COMMENT

96026280
Contact: Kerlavage, AR
Bioinformatics
The Institute for Genomic Research
9712 Medical Center Drive, Rockville, MD 20850 USA
Tel: 3018699056
Fax: 3018699423

Email: arkerlavage@tigr.org
For clone availability, additional sequence and expression information related to this EST, please check the TIGR Human Gene Index (http://www.tigr.org/tdb/hgi/hgi.html)
Seq primer: M13 Reverse.

FEATURES

Source

1..337

/organism="Homo sapiens"
/db_xref="ATCC (inhost):167143"
/db_xref="taxon:9606"

/clone_lib="Ovary II"

/sex="Female"

/dev_stage="adult"
/note="Organ: ovary; Vector: pbluescript SK-; Site_1: EcoRI; Site_2: XhoI"

BASE COUNT 86 a 72 c 82 g 94 t 3 others

Query Match 85.5%; Score 81.2; DB 9; Length 337;
Best Local Similarity 90.5%; Pred. No. 3.5e-16;
Matches 86; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 1 CCGCTTATCGCCAGCTCTCTCAGAGAAAGAGAGAGGCGCATTCCTGAGAGAGA 60
|||||
Db 184 CCGCTTATCGACTAAGCTCTCTCAGAGAAAGAGAGAGGCGCATTCAGAGAGAGA 125
QY 61 CTGGCAACTGATTTTACCCCAAGCCCAACCTCA 95
|||||
Db 124 CTGGCACTAGATTATTCACCATGCCCAAAATCTCA 90
RESULT 6
HSA0408776 738 bp DNA linear GSS 02-AUG-2001
LOCUS Homo sapiens GSS, clone BAC RPCI-11 447F21 Sp6, genomic survey
DEFINITION sequence.
ACCESSION AJ408776
VERSION AJ408776.1 GI:15131033
KEYWORDS GSS; genomic survey sequence.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
REFERENCE
AUTHORS Beyer, K.S., Klauk, S.M., Wiemann, S. and Poustka, A.
TITLE Construction of a physical map of an autism susceptibility region
JOURNAL Gene 272 (1-2), 85-91 (2001)
MEDLINE 21363428
PUBMED 11470513
REFERENCE
AUTHORS Beyer, K.S.
TITLE Direct Submission
SUBMITTED (12-FEB-2001) Beyer K.S., Molecular Genome Analysis,
Deutsches Krebsforschungszentrum, Im Neuenheimer Feld 280,
Heidelberg, 69120, GERMANY
FEATURES
source location/Qualifiers
1..738
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="7"
/map="7q32.3-33"
/clone="BAC RPCI-11 447F21 Sp6"
/sex="male"
/clone_1lb="RPCI-11"
BASE COUNT 241 a 174 c 160 g 163 t
ORIGIN
Query Match 81.7%; Score 77.6; DB 17; Length 738;
Best Local Similarity 94.8%; Pred. No. 6.8e-15;
Matches 91; Conservative 0; Mismatches 4; Indels 1; Gaps 1;
QY 1 CCGCTTATCGCCAGCTCTCTCAGAGAAAGAGAGGCGCATTCCTGAGAGAGA 60
|||||
Db 507 CCGCTTATCGCCAGCTCTCTCAGAGAAAGAGAGGCGCATTCAGAGAGAGA 566
QY 61 CTGGCAACTGA-TTTACCCCAAGCCCAACCTCA 95
|||||
Db 567 CTGGCACTAGATTATTCACCATGCCCAAAATCTCA 602
RESULT 7
B1772113 723 bp mRNA linear EST 25-SEP-2001
LOCUS B1772113
DEFINITION 603059196F1 NIH_MGC_122 Homo sapiens CDNA IMAGE:5208814 5',
mRNA sequence.
ACCESSION B1772113
VERSION B1772113.1 GI:15763691
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
REFERENCE
1 (bases 1 to 723)

AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: Life Technologies, Inc.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: L1AM1523 row: 0 column: 23
High quality sequence stop: 722.
FEATURES
source location/Qualifiers
1..723
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5208814"
/clone_1lb="NIH_MGC_122"
/lab_host="DH10B"
/note="Organ: pooled lung and spleen; Vector: pCMV-Sport6;
Site_1: Not; Site_2: EcoRV (destroyed); RNA source
anonymous pool of 24 week female lung, 16 week female
spleen, and 20-22 week male spleens. Library is oligo-dT
primed and directionally cloned (EcoRV site is destroyed
upon cloning). Average insert size 1.4 kb, insert size
range 1-3 kb. Library is normalized and enriched for
full-length clones and was constructed by C. Gruber
(Invitrogen). Research Genetics tracking code 026. Note:
this is a NIH-MGC Library."
BASE COUNT 229 a 185 c 141 g 168 t
ORIGIN
Query Match 73.7%; Score 70; DB 13; Length 723;
Best Local Similarity 84.0%; Pred. No. 2.1e-12;
Matches 79; Conservative 0; Mismatches 15; Indels 0; Gaps 0;
QY 1 CCGCTTATCGCCAGCTCTCTCAGAGAAAGAGAGGCGCATTCCTGAGAGAGA 60
|||||
Db 121 CCGCTTATCGCCAGCTCTCTCAGAGAAAGAGAGGCGCATTCAGAGAGAGA 180
QY 61 CTGGCAACTGATTTTACCCCAAGCCCAACCTC 94
|||||
Db 181 CTGGCACTAGATTATTCACCATGCCCAAAATCTCA 214
RESULT 8
BF761543 421 bp mRNA linear EST 12-JAN-2001
LOCUS BF761543
DEFINITION CM0-CS0041-051000-592-c12 CS0041 Homo sapiens CDNA, mRNA sequence.
ACCESSION BF761543
VERSION BF761543.1 GI:12109443
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
REFERENCE
1 (bases 1 to 421)
Dias Neto, E., Garcia Correa, R., Verjowski-Almeida, S., Briones, M.R.,
Nagai, M.A., da Silva, M. Jr., Zago, M.A., Bordin, S., Costa, F.F.,
Goldman, G.H., Carvalho, A.F., Matsukuma, A., Baia, G.S., Simpson, D.H.,
Brunstein, A., de Oliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare,
M.J., Soares, F., Brentani, R.S., Reis, L.F., de Souza, S.J. and
Simpson, A.J.
Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
JOURNAL
MEDLINE
COMMENT Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,

Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?tbl=CM0&ct=CM0-CS0041-051000-592-cl2&ts=2000-10-05&ct4=1)
Seq primer: puc 18 forward
High quality sequence start: 11
High quality sequence stop: 421.
Location/Qualifiers
1. 421
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="CS0041"
/note="Organ: colon_est; Vector: puc18; Site_1: SmaI; Site_2: SmaI; A mini-library was made by cloning products derived from ORESRES PCR (U.S. letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."
BASE COUNT 122 a 111 c 86 g 102 t
ORIGIN
Query Match 71.4%; Score 67.8; DB 12; Length 421;
Best Local Similarity 82.1%; Pred. No. 9.2e-12;
Matches 78; Conservative 0; Mismatches 17; Indels 0; Gaps 0;
QY 1 CCGCCTTATCGCAGCTCTTCAGAGAAAGAACAGCCATTACCTGGAGAGA 60
DB 258 CCGCGCTTATCGCAGCTCTTCAGAGAAAGAACAGCCATTACCGAGAGACAC 317
QY 61 CTGGCAACTGATTTTACCACCAAGCCCAACTCA 95
DB 318 TGGCACTAGATTATTACCACCAAGCCCAACTCA 352
RESULT 9
LOCUS AAI79079 352 bp mRNA linear EST 31-DEC-1996
DEFINITION zplie01.r1 Stragene fetal retina 937202 Homo sapiens cDNA clone IMAGE:609144 5' similar to SW:POL_AVIRE P03360 POL POLYPROTEIN ;,
ACCESSION AAI79079
VERSION AAI79079.1 GI:1760431
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 352)
Hillier,L., Lennon,G., Becker,M., Bonaldo,M.F., Chiapelli,B., Chisoe,S., Dietrich,N., Dubuque,T., Favello,A., Gish,W., Hawkins,M., Hultman,M., Kucaba,T., Lacy,M., Le,M., Le,N., Marids,E., Moore,B., Morris,M., Parsons,J., Prange,C., Rifkin,L., Rohlfing,T., Schellenberg,K., Soares,M.B., Tan,F., Thierry-Mieg,J., Trevasks,E., Underwood,K., Wohldmann,P., Waterston,R., Wilson,R. and Marra,M. Generation and analysis of 280,000 human expressed sequence tags Genome Res. 6 (9), 807-828 (1996)
97044478
TITLE Contact: Wilson RK
JOURNAL Washington University School of Medicine
MEDLINE 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
COMMENT Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available royalty-free through LNLN ; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
Seq primer: -28M13 rev2 from Amersham
High quality sequence stop: 320.

FEATURES
source Location/Qualifiers
1. 352
/organism="Homo sapiens"
/db_xref="GDB:4624403"
/db_xref="taxon:9606"
/clone_lib="IMAGE:609144"
/clone_lib="Stratagene fetal retina 937202"
/sex="mixed"
/lab_host="SOLR (kanamycin resistant)"
/note="Vector: pbluescript SK-; Site_1: EcoRI; Site_2: XhoI; Cloned unidirectionally. Primer: Oligo dT. Pooled retinal tissue. Average insert size: 1.0 kb; Uni-ZAP XR Vector: -5' adaptor sequence: 5' GATTGGGCGACGAG 3' -3' adaptor sequence: 5' CTCGAGTTTCTTTTCTTTTCTTTT 3' "

BASE COUNT 114 a 88 c 65 g 80 t 5 others
ORIGIN
Query Match 62.1%; Score 59; DB 9; Length 352;
Best Local Similarity 91.3%; Pred. No. 6.6e-09;
Matches 84; Conservative 0; Mismatches 6; Indels 2; Gaps 2;
QY 1 CCGCCTTATCGCAGCTCTTCAGAGAAAGAACAGGCCA-TTACCTGGAGAG 59
DB 247 CCGCCTTATCGCAGCTCTTCAGAGAAAGAACAGGCCATTACCGAGAGAG 306
QY 60 ACTGGCACT-GATTTTACCACAGCCCAA 90
DB 307 ACTGGCACTAGATTTTACCCACATGCGCAA 338
RESULT 10
LOCUS AAI76759/c 508 bp mRNA linear EST 30-DEC-1996
DEFINITION zplie01.s1 Stragene fetal retina 937202 Homo sapiens cDNA clone IMAGE:609144 3' similar to SW:POL_MLVK P31795 POL POLYPROTEIN ;,
ACCESSION AAI76759
VERSION AAI76759.1 GI:1758064
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 508)
Hillier,L., Lennon,G., Becker,M., Bonaldo,M.F., Chiapelli,B., Chisoe,S., Dietrich,N., Dubuque,T., Favello,A., Gish,W., Hawkins,M., Hultman,M., Kucaba,T., Lacy,M., Le,M., Le,N., Marids,E., Moore,B., Morris,M., Parsons,J., Prange,C., Rifkin,L., Rohlfing,T., Schellenberg,K., Soares,M.B., Tan,F., Thierry-Mieg,J., Trevasks,E., Underwood,K., Wohldmann,P., Waterston,R., Wilson,R. and Marra,M. Generation and analysis of 280,000 human expressed sequence tags Genome Res. 6 (9), 807-828 (1996)
97044478
TITLE Contact: Wilson RK
JOURNAL Washington University School of Medicine
MEDLINE 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
COMMENT Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available royalty-free through LNLN ; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
Seq primer: -40M13 fwd. from Amersham
High quality sequence stop: 270.
Location/Qualifiers
1. 508
/organism="Homo sapiens"
/db_xref="GDB:4624403"
/db_xref="taxon:9606"
/clone_lib="IMAGE:609144"
/clone_lib="Stratagene fetal retina 937202"
/sex="mixed"
/lab_host="SOLR (kanamycin resistant)"
/note="Vector: pbluescript SK-; Site_1: EcoRI; Site_2:

XhoI: Cloned unidirectionally. Primer: Oligo dT. Pooled retinal tissue. Average insert size: 1.0 kb; Uni-ZAP XR Vector: -5' adaptor sequence: 5' GATTCCGCGACGAG 3' -3' adaptor sequence: 5' CTCGAGTTTCTTTTCTTTT 3' "

BASE COUNT 112 a 99 c 130 g 163 t 4 others

Query Match 58.5%; Score 55.6; DB 9; Length 508;
Best Local Similarity 86.7%; Pred. No. 9,7e-08;
Matches 72; Conservative 0; Mismatches 10; Indels 1; Gaps 1;

QY 1 CCTGCTTATCGCCAAAGCTCTTCAGAGAA-CAAGAAGACGCGATTACCTCGAGAG 59
Db 482 CTTCCTTATTCACCAAGCCTTCAGAGAACCAAGAGCGCATTAACCGAGAG 423

QY 60 ACTGCACACTGATTTTACCACA 82
Db 422 ACTGCACACTGATTTTACCACA 400

RESULT 11
A0469362/c
LOCUS A0469362 660 bp DNA linear GSS 23-APR-1998

DEFINITION C17B1-E1-2595M14.TR C17B1-E1 Homo sapiens genomic clone 2595M14,
DNA sequence.

ACCESSION A0469362
VERSION A0469362.1 GI:4653252

KEYWORDS GSS.

SOURCE human.

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 660)
Zhang, S., Adams, M.D., Nierman, W., Malek, J., Shizuya, H., Simon, M. and
Venter, J.C.

TITLE Use of BAC End Sequences from Caltech Libraries for Sequence-Ready
Map Building

UNPUBLISHED (1997)

Other_GSSs: C17B1-E1-2595M14.TF

Contact: Shaying Zhao, William Nierman, Mark Adams

Department of Eukaryotic Genomics

The Institute for Genomic Research

9712 Medical Center Dr., Rockville, MD 20850

Tel: 301 838 0200

Fax: 301 838 0208

Email: hbeetlgr.org

Clones are available from Research Genetics (Info@resgen.com). BAC
end search page:

http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html.

Seq primer: M13 Reverse

Class: BAC ends.

FEATURES
Location/Qualifiers

1..660
/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="2595M14"

/clone_lib="C17B1-E1"

/sex="male"

/cell_type="sperm"

/note="Vector: pBelobAC11; Site_1: EcoRI; Site_2: EcoRI;
Caltech Human BAC Library D"

BASE COUNT 158 a 137 c 146 g 219 t

ORIGIN

Query Match 44.6%; Score 42.4; DB 17; Length 660;
Best Local Similarity 72.4%; Pred. No. 0.0022;
Matches 55; Conservative 0; Mismatches 21; Indels 0; Gaps 0;

QY 15 AAGCTCTTCAGAGAGAAAGAGAGGCGCATTAACCTCGAGAGAGCTGCACTGATT 74
Db 369 AGGCTCTTCAGAGAGAGAGAGAGGCGCATTAACCTCGAGAGAGCTGCACTGATT 310

QY 75 TACCCACAAGCCCAA 90

Db 309 TTACCCATATGCCAA 294

RESULT 12
A0109046/c
LOCUS A0109046 385 bp DNA linear GSS 29-AUG-1998

DEFINITION C17-HSP-2377A19.TR C17-HSP Homo sapiens genomic clone 2377A19, DNA
sequence.

ACCESSION A0109046
VERSION A0109046.1 GI:3485736

KEYWORDS GSS.

SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 385)
Adams, M.D., Rounsley, S.D., Zhao, S., Bass, S., Linher, K., Golden, K.,
Berry, K., Granger, D., Suh, E., Wible, C., Shizuya, H., Simon, M. and
Venter, J.C.

AUTHORS

TITLE Use of a random human BAC End Sequence Database for Sequence-Ready
Map Building

UNPUBLISHED (1998)

Other_GSSs: C17-HSP-2377A19.TF

Contact: Mark Adams

Department of Eukaryotic Genomics

The Institute for Genomic Research

9712 Medical Center Dr., Rockville, MD 20850, USA

Tel: 301 838 0200

Fax: 301 838 0208

Email: mdadams@tigr.org

Clones are available from Research Genetics (Info@resgen.com). BAC
end search page:

http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html.

Seq primer: M13 Reverse

Class: BAC ends.

FEATURES
Location/Qualifiers

1..385
/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="2377A19"

/clone_lib="C17-HSP"

/sex="Male"

/cell_type="Sperm"

/note="Vector: pBelobAC11; Site_1: HindIII; Site_2:
HindIII"

BASE COUNT 100 a 77 c 88 g 120 t

ORIGIN

Query Match 39.2%; Score 37.2; DB 17; Length 385;
Best Local Similarity 65.9%; Pred. No. 0.094;
Matches 54; Conservative 0; Mismatches 28; Indels 0; Gaps 0;

QY 9 ATCGCCAAAGCTCTTCAGAGAGAAAGAGGCGCATTAACCTCGAGAGAGCTGCAAC 68
Db 223 ATCATTAAGGCGCCCTTGGGGAACAAGATGACTACCGAGAGAGCTGCGAGT 164

QY 69 TGATTTTACCCACAAGCCCAA 90
Db 163 TAGACTTCACCCATATGCTTA 142

RESULT 13
A0196669/c
LOCUS A0196669 439 bp DNA linear GSS 16-SEP-1998

DEFINITION C17-HSP-2381B2.TR C17-HSP Homo sapiens genomic clone 2381B2, DNA
sequence.

ACCESSION A0196669
VERSION A0196669.1 GI:3604031

KEYWORDS GSS.

SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

REFERENCE	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
AUTHORS	Adams,M.D., Rounsley,S.D., Zhao,S., Bass,S., Linher,K., Golden,K., Bentley,K.K., Granger,D., Suh,E., Wildie,C., Shizuya,H., Simon,M. and Venier,J.C.
TITLE	Use of a random human BAC End Sequence Database for Sequence-Ready Map Building
JOURNAL	Unpublished (1998)
COMMENT	Other GSSS: CIT-HSP-2381B2.TF Contact: Mark Adams Department of Eukaryotic Genomics The Institute for Genomic Research 9712 Medical Center Dr., Rockville, MD 20850, USA Tel.: 301 838 0200 Fax: 301 838 0208 Email: mdadams@tigr.org Clones are available from Research Genetics (info@resgen.com). BAC end search page: http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html . Seq primer: M13 Reverse Class: BAC ends.
FEATURES	Location/Qualifiers
SOURCE	1..439
	/organism="Homo sapiens"
	/db_xref="taxon:9606"
	/clone="2381B2"
	/clone_1lb="CIT-HSP"
	/sex="Male"
	/cell_type="Sperm"
	/note="Vector: pBelobAC11; site_1: HindIII; site_2:
BASE COUNT	HindIII"
ORIGIN	109 a 105 c 97 g 128 t
Query Match	39.2%; Score 37.2; DB I7; Length 439;
Best Local Similarity	63.3%; Pred. No. 0.098;
Matches	57; Conservative 0; Mismatches 33; Indels 0; Gaps 0;
OY	1 CCTGCTTATGTGCCAAGTCCTTCAGAGAACAAGAACAGCCCATTTACCCTGGAGANGA 60
Db	157 CTGTGTATATCATTAAGCGCCCTGTGGGGGAACAAGATAGTCACTATCCCGGAGAGA 98
OY	61 CTGGCAACTGATTTTTACCACACAGCCCAA 90
Db	97 CTGGCAGTTTAGCTTCTCCCATATGCCCTAA 68
RESULT 14	
BM713850	604 bp mRNA linear EST 28-FEB-2002
LOCUS	U1-E-EJ0-ahp-j-24-0-U1.r1 U1-E-EJ0 Homo sapiens cDNA clone
DEFINITION	U1-E-EJ0-ahp-j-24-0-U1 5', mRNA sequence.
ACCESSION	BM713850
VERSION	BM713850.1 GI:19027108
KEYWORDS	EST.
SOURCE	human.
ORGANISM	Homo sapiens
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo. 1 (bases 1 to 604) Bonaldio,M.F., Lennon,G. and Soares,M.B. Normalization and subtraction: two approaches to facilitate gene discovery Genome Res. 6 (9), 791-806 (1996)
JOURNAL	97044477
MEDLINE	Contact: Soares, MB
COMMENT	Program for Rat Gene Discovery and Mapping University of Iowa 451 Eckstein Medical Research Building Iowa City, IA 52242, USA Tel.: 319 335 8250 Fax: 319 335 9565 Email: mscares@blue.weeg.iowa.edu Tissue Procurement: Dr. Gregg Hageman

FEATURES	SOURCE
CDNA Library preparation: Dr. M. Bento Soares, University of Iowa CDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa DNA Sequencing by: Dr. M. Bento Soares, University of Iowa Clone Distribution: Researchers may obtain clones from Research Genetics (www.resgen.com).	
The following repetitive elements were found in this CDNA sequence: 2.604, >HERV9#LTR/Retroviral	
Seq primer: M13 Reverse.	
Location/Qualifiers	
1. 604	
/organism="Homo sapiens"	
/db_xref="taxon:9606"	
/clone="UI-E-EJ0-ahp-j-24-0-UI"	
/clone_11b="UI-E-EJ0"	
/tissue_type="fetal eyes, lens, eye anterior segment, optic nerve, retina, Retina Foveal and Macular, RPE and Choroid"	
/dev_stage="fetal and adult"	
/lab_host="DH10B (life technologies) (T1 phage resistant)"	
/note="Organ: eye; Vector: pRT73-Pac (Pharmacia) with a modified polylinker; Site.1: Ecor I; Site.2: Not I; UI-E-EJ0 is a subtracted cDNA library constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was ligated to an Ecor I adaptor, digested with Not I, and cloned directionally into pRT73-Pac vector. The oligonucleotide used to prime the synthesis of first-strand cDNA contains a library tag sequence that is located between the Not I site and the (d)18 tail. The sequence tags for this library are: Fetal eyes, AGATCATACA; lens, CGATTAGCGA; eye anterior segment, AATGCGGAT; optic nerve, CCATTAGTGT; retina, CCGCG; Retina Foveal and Macular, GTCC; RPE and Choroid, ACCCTA. This library was created for the program, Gene Discovery in the Visual System, supported by National Eye Institute (NEI)."	
BASE COUNT	183 a 139 c 134 g 148 t
ORIGIN	
Query Match	39.2%; Score 37.2; DB 14; Length 604;
Best Local Similarity	65.9%; Pred. No. 0.11;
Matches 54; Conservative 0; Mismatches 28; Indels 0; Gaps 0;	
QY	9 ATCCGCAAGCTCTTCAGSAGACAAGAAGAGCCATTACCTCGAGAGACTGGCCAC 68
Db	316 ATCATPAGGCCCCCTTGGGGGACACAAAGATGCTACTATCTTGGAGAGACTGGCGCT 375
QY	69 TGATTTTACCACAAGCCCAAA 90
Db	376 TAGACTTACCCTCATATGCTTAA 397
RESULT 15	
LOCUS	B93732 489 bp DNA linear GSS 25-JUN-1998
DEFINITION	CIT-HSP-2165119, TR CIT-HSP Homo sapiens genomic clone 2165119, DNA
VERSION	B93732
KEYWORDS	B93732.1 GI:2976069
SOURCE	GSS.
ORGANISM	human.
	Homo sapiens
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE	1 (bases 1 to 489)
AUTHORS	Adams,M.D., Rounsley,S.D., Zhao,S., Field,C.E., Bass,S., Linher,K., Golden,K., Berry,K., Granger,D., Suh,E., Wible,C., Shizuya,H., Simon,M. and Verter,J.C.
TITLE	Use of a random BAC End Sequence Database for Sequence-Ready Map Building (1998)
JOURNAL	Unpublished (1998)
COMMENT	Other:GSS: CIT-HSP-2165119, TF
	Contact: Mark Adams

Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208

Email: mdadams@tigr.org
Clones are available from Research Genetics (Info@resgen.com). BAC
end search page:
http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html
Seq primer: M13 Reverse
Class: BAC ends.

FEATURES

SOURCE

Location/Qualifiers
1..489

/organism="Homo sapiens"
/db_xref="GDB:7101197"
/db_xref="taxon:9606"
/clone="2165119"
/clone_11b="CIT-HSP"
/sex="Male"
/cell_type="Sperm"
/note="Vector: pBelobAC11; Site_1: HindIII; Site_2:
HindIII"

BASE COUNT 160 a 112 c 100 g 117 t
ORIGIN

Query Match

37.1%; Score 35.2; DB 17; Length 489;

Best Local Similarity 68.1%; Pred. No. 0.46; Mismatches 23; Indels 0; Gaps 0;

Matches 49; Conservative 0; Mismatches 23; Indels 0; Gaps 0;

OY 19 TCCTTCAGGAGAACAAAGACAGCCATTACCTGGAGAGAGACTGGCAACTGATTTTACC 78
111 111111 11111111 1111 11 111111 111111 111 1
Db 167 TCCTTCAGGAGATCCAAAGAGACAGAGCTACTCTGGGAAGAGACTAGCAATGATTTTCAC 226
OY 79 CACAGCCCAAA 90
111 11111111
Db 227 CCTATGCCAAA 238

Search completed: May 2, 2003, 12:25:23
Job time : 1098.14 secs

GenCore version 5.1.5
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 2, 2003, 11:30:30 ; Search time 38 Seconds
(without alignments)
766.692 Million cell updates/sec

Title: US-09-719-554-3_COPY_6956_7050

Perfect score: 95
Sequence: 1 cctgccttatgcgaagctc.....accacaagccaactca 95

Scoring table: IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Searched: 441362 seqs, 15338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : Issued_Patents_MA:*
1: /cgn2_6/ptodata/2/ina/5A.COMB.seq:*
2: /cgn2_6/ptodata/2/ina/5B.COMB.seq:*
3: /cgn2_6/ptodata/2/ina/6A.COMB.seq:*
4: /cgn2_6/ptodata/2/ina/6B.COMB.seq:*
5: /cgn2_6/ptodata/2/ina/PCTUS.COMB.seq:*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	94	98.9	2946	4 US-09-175-928-3	Sequence 3, Appli
2	90.4	95.2	374	1 US-08-686-878A-47	Sequence 47, Appli
3	90.4	95.2	374	1 US-08-721-489-1	Sequence 1, Appli
4	27	28.4	34185	4 US-09-545-481-3	Sequence 31, Appli
5	26.2	27.6	3234	1 US-08-264-534-31	Sequence 31, Appli
6	26.2	27.6	3234	1 US-08-083-590A-10	Sequence 10, Appli
7	26.2	27.6	3234	1 US-08-465-500-31	Sequence 31, Appli
8	26.2	27.6	3234	2 US-08-346-128-31	Sequence 31, Appli
9	26.2	27.6	3234	3 US-08-532-384-10	Sequence 10, Appli
10	26.2	27.6	3234	3 US-08-893-828-31	Sequence 31, Appli
11	26.2	27.6	4131	1 US-08-309-512-3	Sequence 3, Appli
12	26.2	27.6	4131	5 PCT-US92-08756A-3	Sequence 3, Appli
13	25.8	27.2	3002	4 US-08-369-364A-1	Sequence 1, Appli
14	25.4	26.7	2707	4 US-08-992-035A-5	Sequence 5, Appli
15	25.2	26.5	52297	4 US-09-426-436-1	Sequence 1, Appli
16	25.2	26.5	52297	4 US-08-705-557-1	Sequence 585, App
17	25	26.3	588	4 US-09-221-017B-585	Sequence 585, App
18	25	26.3	3789	1 US-07-872-644-42	Sequence 42, Appli
19	25	26.3	3789	1 US-08-297-494-42	Sequence 42, Appli
20	25	26.3	3789	1 US-08-297-510-42	Sequence 42, Appli
21	25	26.3	3789	1 US-08-479-532-42	Sequence 42, Appli
22	25	26.3	3789	1 US-08-455-526-42	Sequence 42, Appli
23	25	26.3	3789	1 US-08-455-525-42	Sequence 42, Appli
24	25	26.3	3789	3 US-09-139-491-42	Sequence 42, Appli
25	25	26.3	3789	5 PCT-US92-03222-42	Sequence 42, Appli
26	25	26.3	4131	1 US-07-872-644-38	Sequence 38, Appli
27	25	26.3	4131	1 US-08-297-494-38	Sequence 38, Appli

28	25	26.3	4131	1 US-08-297-510-38	Sequence 38, Appli
29	25	26.3	4131	1 US-08-479-532-38	Sequence 38, Appli
30	25	26.3	4131	1 US-08-455-526-38	Sequence 38, Appli
31	25	26.3	4131	1 US-08-455-525-38	Sequence 38, Appli
32	25	26.3	4131	3 US-09-139-491-38	Sequence 38, Appli
33	25	26.3	4131	5 PCT-US92-03222-38	Sequence 38, Appli
34	24.8	26.1	1574	4 US-09-173-581-12	Sequence 12, Appli
35	24.8	26.1	1574	4 US-09-420-915-12	Sequence 12, Appli
36	24.8	26.1	4661	4 US-09-221-017B-970	Sequence 970, App
37	24.6	25.9	314	4 US-09-328-111-35	Sequence 35, Appli
38	24.6	25.9	733	4 US-08-998-416-989	Sequence 989, App
39	24.6	25.9	4558	1 US-08-309-512-2	Sequence 2, Appli
40	24.6	25.9	4558	5 PCT-US92-08756A-2	Sequence 2, Appli
41	24.2	25.5	1058	4 US-08-818-112-45	Sequence 45, Appli
42	24.2	25.5	1058	4 US-08-818-111-45	Sequence 45, Appli
43	24.2	25.5	1058	4 US-09-036-536-45	Sequence 45, Appli
44	24.2	25.5	1058	4 US-09-072-596-45	Sequence 45, Appli
45	24.2	25.5	1620	1 US-07-616-022C-9	Sequence 9, Appli

ALIGNMENTS

```
RESULT 1
US-09-175-928-3
: Sequence 3, Application US/09175928A
: Patent No. 6312921
: GENERAL INFORMATION:
: APPLICANT: Jacobs, Kenneth
: APPLICANT: McCoy, John M.
: APPLICANT: Lavalley, Edward R.
: APPLICANT: Cavallie-Racie, Lisa A.
: APPLICANT: Evans, Cheryl
: APPLICANT: Merderg, David
: APPLICANT: Treacy, Maurice
: APPLICANT: Mi, Sha
: APPLICANT: Genetics Institute, Inc.
: TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES ENCODING THEM
: FILE REFERENCE: 6006B, A1172A
: CURRENT APPLICATION NUMBER: US/09/175, 928A
: CURRENT FILING DATE: 1998-10-20
: NUMBER OF SEQ ID NOS: 62
: SOFTWARE: Patentln Ver. 2.0
: SEQ ID NO 3
: LENGTH: 2946
: TYPE: DNA
: ORGANISM: Homo sapiens
US-09-175-928-3
Query Match          98.9%  Score 94:  DB 4:  Length 2946;
Best Local Similarity 100.0%:  Pred. No. 4.2e-25;
Matches 94:  Conservative 0;  Mismatches 0;  Indels 0;  Gaps 0;
OY      2  CTGCTTATGCGCAAGCTCTTCAGAGACAAAGACAGGCATTACCTGAGAGAC 61
DB      6  CTGCTTATGCGCAAGCTCTTCAGAGACAAAGACAGGCATTACCTGAGAGAC 65
OY      62  TGGCACTGATTTTACCCCAAGCCCAACTCA 95
DB      66  TGGCACTGATTTTACCCCAAGCCCAACTCA 99
RESULT 2
US-08-686-878A-47
: Sequence 47, Application US/08686878A
: Patent No. 5708157
: GENERAL INFORMATION:
: APPLICANT: Jacobs, Kenneth
: APPLICANT: McCoy, John
: APPLICANT: Lavalley, Edward
: APPLICANT: Racie, Lisa
: APPLICANT: Merderg, David
: APPLICANT: Treacy, Maurice
```

```

; APPLICANT: Evans, Cheryl
; APPLICANT: Spaulding, Ylki
; TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES
; TITLE OF INVENTION: SECRETED THEM
; NUMBER OF SEQUENCES: 71
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genetics Institute, Inc.
; STREET: 87 CambridgePark Drive
; CITY: Cambridge
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02140
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/686,878A
; FILING DATE:
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Brown, Scott A.
; REGISTRATION NUMBER: 32,724
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 498-8224
; TELEFAX: (617) 876-5851
; INFORMATION FOR SEQ ID NO: 47:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 374 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; US-08-686-878A-47

Query Match          95.2%; Score 90.4; DB 1; Length 374;
Best Local Similarity 96.8%; Pred. No. 3.9e-24;
Matches 91; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 CTCGCTTATGCGCAAGCTCTTCAGAGAAACAAGAGCCATTACCTGGAGAAGC 61
DB 29 CTCGCTTATGCGCAAGCTCTTCAGAGAAACAAGAGCCATTACCTGGAAAAA 88
QY 62 TGGCAACTGATTTTACCCACAAGCCCAACCTCA 95
DB 89 TGGCAACTGATTTTACCCACAAGCCCAACCTCA 122

RESULT 3
US-08-721-489-1
; Sequence 1, Application US/08721489
; Patent No. 5786465
; GENERAL INFORMATION:
; APPLICANT: Jacobs, Kenneth
; APPLICANT: McCoy, John
; APPLICANT: Lavallee, Edward
; APPLICANT: Racie, Lisa
; APPLICANT: Merberg, David
; APPLICANT: Treacy, Maurice
; TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES
; TITLE OF INVENTION: ENCODING THEM
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genetics Institute, Inc.
; STREET: 87 CambridgePark Drive
; CITY: Cambridge
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02140
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
```

```

; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/721,489
; FILING DATE:
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Brown, Scott A.
; REGISTRATION NUMBER: 32,724
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 498-8224
; TELEFAX: (617) 876-5851
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 374 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; US-08-721-489-1

Query Match          95.2%; Score 90.4; DB 1; Length 374;
Best Local Similarity 96.8%; Pred. No. 3.9e-24;
Matches 91; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 CTCGCTTATGCGCAAGCTCTTCAGAGAAACAAGAGCCATTACCTGGAGAAGC 61
DB 29 CTCGCTTATGCGCAAGCTCTTCAGAGAAACAAGAGCCATTACCTGGAAAAA 88
QY 62 TGGCAACTGATTTTACCCACAAGCCCAACCTCA 95
DB 89 TGGCAACTGATTTTACCCACAAGCCCAACCTCA 122

RESULT 4
US-09-545-481-3
; Sequence 3, Application US/09545481
; Patent No. 6451319
; GENERAL INFORMATION:
; APPLICANT: Chiang, Christina H.
; APPLICANT: Cochran, Mark D.
; TITLE OF INVENTION: NO. 6451319el Recombinant And Mutant Adenoviruses
; FILE REFERENCE: SY0993K US
; CURRENT APPLICATION NUMBER: US/09/545,481
; CURRENT FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: 60/128,766
; PRIOR FILING DATE: 1999-04-09
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 3
; LENGTH: 34185
; TYPE: DNA
; ORGANISM: Bovine adenovirus type-1
; US-09-545-481-3

Query Match          28.4%; Score 27; DB 4; Length 34185;
Best Local Similarity 56.0%; Pred. No. 4.6;
Matches 51; Conservative 0; Mismatches 40; Indels 0; Gaps 0;

QY 1 CTCGCTTATGCGCAAGCTCTTCAGAGAAACAAGAGCCATTACCTGGAGAAGA 60
DB 24080 CCGCCATTCTCGGCCAATATACATGATATAAAAAAGCGGGAATAATCTTCTTAAAA 24139
QY 61 CTGGCAACTGATTTTACCCACAAGCCCAAC 91
DB 24140 AGGGCCACGCGGTACTAGACCCCAAAAC 24170

RESULT 5
US-08-264-534-31
; Sequence 31, Application US/08264534
; Patent No. 5648464
```



```

CORRESPONDENCE ADDRESS:
ADDRESSEE: PENNIE & EDMONDS
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: NY
COUNTRY: USA
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/893,828
FILING DATE: 11-JUL-1997
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Mistrock, S. Leslie
REGISTRATION NUMBER: 18,872
REFERENCE/DOCKET NUMBER: 7326-050
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-8864/9741
INFORMATION FOR SEQ ID NO: 31:
SEQUENCE CHARACTERISTICS:
LENGTH: 3234 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: unknown
MOLECULE TYPE: CDNA
US-08-893-828-31

```

```

Query Match          27.6%:  Score 26.2;  DB 3;  Length 3234;
Best Local Similarity 67.3%:  Pred. No.3.5;
Matches 37;  Conservative 0;  Mismatches 18;  Indels 0;  Gaps 0;

QY      11  CGCCAAGTCCTTCAGAGAACAAAGACAGGCCATTACCCCTGGAGAGACTGGC 65
          |||||  |||  |||  ||  |||  ||  |||  ||  |||  ||  |||
Db      1470  CGCCAAAGGCGCTGCTGGAGGCCAGCGGCAAGATCCCAACATCCAGAGAACACATGGCC 1524

RESULT 11
US-08-309-512-3
: Sequence 3, Application US/08309512
: Patent No. 5759828
: GENERAL INFORMATION:
: APPLICANT: Tal, Ronny
: APPLICANT: Benziman, Moshe
: APPLICANT: Gelfand, David H.
: APPLICANT: Ben-Bassat, Arle
: APPLICANT: Calhoon, Roger D.
: APPLICANT: Wong, Hing C.
: TITLE OF INVENTION: CYCLIC DIGUANYLATE METABOLIC ENZYMES
: NUMBER OF SEQUENCES: 63
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Pennie & Edmonds
: STREET: 2730 Sand Hill Road
: CITY: Menlo Park
: STATE: California
: COUNTRY: U.S.A.
: ZIP: 94025
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patentin Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/309,512
: FILING DATE:
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 07/800,218
: FILING DATE: 29-NOV-1991

```

```

1  ATTORNEY/AGENT INFORMATION:
2  NAME:  Bortner, Scott R.
3  REGISTRATION NUMBER:  34,296
4  REFERENCE/DOCKET NUMBER:  8145-008
5  TELECOMMUNICATION INFORMATION:
6  TELEPHONE:  (415) 854-3660
7  TELEFAX:  (415) 854-3694
8  TELEEX:  66141 PENNIE
9  INFORMATION FOR SEQ ID NO:  3:
10 SEQUENCE CHARACTERISTICS:
11     LENGTH:  4131 base pairs
12     TYPE:  nucleic acid
13     STRANDEDNESS:  double
14     TOPOLOGY:  linear
15     MOLECULE TYPE:  DNA (genomic)
16     HYPOTHETICAL:  NO
17     ANTI-SENSE:  NO
18     ORIGINAL SOURCE:
19     ORGANISM:  Acetobacter xylinum
20
21 US-08-309-512-3

```

	Query Match	27.6%	Score 26.2;	DB 1;	Length 4131;	
	Best Local Similarity	60.6%;	Pred. No. 3.9;	Mismatches 28;	Indels 0;	Gaps 0;
	Matches 43;	Conservative	0;			
OY	2 CTGCCTTATTCGCCAAGCTCCTTCACGAGACAACAAACAGGGCAATTACCCTGGACAAGC	61				
Dd	2317 CAGCTTCATCAACGATTTTCGACATGACACCAATCCCGAGCGCGTGACCATGGCGGTGAT	2376				
OY	62 TGGCAACTGAT	72				
Dd	2377 CGGCATCGGCT	2387				

RESULT 12
 PCT-US92-08756A-3
 Sequence 3, Application PC/TUS9208756A
 GENERAL INFORMATION:
 APPLICANT: Tal, Ronny
 APPLICANT: Benzlman, Moshe
 APPLICANT: Gelfand, David H.
 APPLICANT: Ben-Bassat, Arle
 APPLICANT: Calhoun, Roger D.
 APPLICANT: Wong, Hing C.
 TITLE OF INVENTION: CYCLIC DIGUANYLATE
 TITLE OF INVENTION: METABOLIC ENZYMES
 NUMBER OF SEQUENCES: 11
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Limbach and Limbach
 STREET: 2001 Ferry Building
 CITY: San Francisco
 STATE: CA
 COUNTRY: USA
 ZIP: 94111
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0,
 SOFTWARE: Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: PCT/US92/08756A
 FILING DATE: 19921014
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 07/800,218
 FILING DATE: 29-NOV-1991
 ATTORNEY/AGENT INFORMATION:
 NAME: Bortner, Scott R.
 REGISTRATION NUMBER: 34,298
 REFERENCE/DOCKET NUMBER: WEYR 20050 USA
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 415-433-4150

TELEFAX: 415-433-8716
TELEX: 278356
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 4131 base pairs
TYPE: NUCLEIC ACID
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Acetobacter xylinum
PCT-US92-08756A-3

Query Match 27.6%; Score 26.2; DB 5; Length 4131;
Best Local Similarity 60.6%; Pred. No. 3.9;
Matches 43; Conservative 0; Mismatches 28; Indels 0; Gaps 0;

QY 2 CTGCGTTATCGCCAGAGCTCCTTCAGAGAACAAAGAGCCATTACCTGAGAGAGAC 61
DB 2317 CAGCTTCATCAACGATTCGAGCATGACACCAATGCCAGGCCGCTGACCATGGCGTGAT 2376

QY 62 TGGCACTGAT 72
DB 2377 CGGCATCGCT 2387

RESULT 13
US-09-369-364A-1
Sequence 1, Application US/09369364A
Patent No. 6391610
GENERAL INFORMATION:
APPLICANT: Apte, Suneel
APPLICANT: Hurskainen, Tiina L.
APPLICANT: Hirohata, Satoshi
TITLE OF INVENTION: Nucleic Acids Encoding Zinc Metalloproteases
FILE REFERENCE: 26473/4007/10-30-00
CURRENT APPLICATION NUMBER: US/09/369,364A
CURRENT FILING DATE: 1999-08-06
NUMBER OF SEQ ID NOS: 31
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 1
LENGTH: 3002
TYPE: DNA
ORGANISM: mus musculus ADAMTS-5
FEATURE:
NAME/KEY: CDS
LOCATION: (18)..(2810)
US-09-369-364A-1

Query Match 27.2%; Score 25.8; DB 4; Length 3002;
Best Local Similarity 58.4%; Pred. No. 4.8;
Matches 45; Conservative 0; Mismatches 32; Indels 0; Gaps 0;

QY 12 GCCAGCTCCTTCAGGAGAACAAAGAGCCATTACCTGAGAGAGACTGCACTGA 71
DB 990 GACACGAGTGTGAGGAGACAGAAATGCGCCGACACCTCAAGAACTTTGCAATGG 1049

QY 72 TTTTACCACAGCCCA 88
DB 1050 CAGCACCACATTAACCA 1066

RESULT 14
US-08-992-035A-5/C
Sequence 5, Application US/08992035A
Patent No. 6242179
GENERAL INFORMATION:
APPLICANT: Shah, Puryl
APPLICANT: Hillman, Jennifer L.
APPLICANT: Corley, Neil C.
APPLICANT: Lal, Preeti

TITLE OF INVENTION: HUMAN PHOSPHATASES
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Dr.
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible

OPERATING SYSTEM: DOS

SOFTWARE: FastSeq for Windows Version 2.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/992,035A

FILING DATE: December 17, 1997

PRIOR APPLICATION DATA:

APPLICATION NUMBER:

FILING DATE:

ATTORNEY/AGENT INFORMATION:

NAME: Billings, Lucy J.

REGISTRATION NUMBER: 36,749

REFERENCE/DOCKET NUMBER: PF-0433 US

TELECOMMUNICATION INFORMATION:

TELEPHONE: 650-845-4166

TELEFAX: 650-845-4166

INFORMATION FOR SEQ ID NO: 5:

SEQUENCE CHARACTERISTICS:

LENGTH: 2707 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

IMMEDIATE SOURCE:

LIBRARY: COLNOR22

CLONE: 1734452

US-08-992-035A-5

Query Match 26.7%; Score 25.4; DB 4; Length 2707;
Best Local Similarity 61.2%; Pred. No. 6.5;
Matches 41; Conservative 0; Mismatches 26; Indels 0; Gaps 0;

QY 21 CTTGAGAGAACAAAGAGCCATTACCTGAGAGAGCTGCAACTGTTTACCA 80

DB 2553 CTTGGGATCAAAAGACCAAGCCAACTCCTTCAGATGGCTGTGTAACCA 2494

QY 81 CAAGCC 87
DB 2493 CAAGCC 2487

RESULT 15
US-09-426-436-1/C
Sequence 1, Application US/09426436
Patent No. 6225066
GENERAL INFORMATION:
APPLICANT: William R. Jacobs, Jr.
APPLICANT: Barry R. Bloom
APPLICANT: Graham F. Hatfull
TITLE OF INVENTION: MYCOBACTERIAL SPECIES-SPECIFIC
REPORTER MYCOBACTERIOPHAGES
NUMBER OF SEQUENCES: 1
CORRESPONDENCE ADDRESS:
ADDRESSEE: Amster, Rothstein & Ebenstein
STREET: 90 Park Avenue
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10016
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 inch 1.44 Mb storage diskette
COMPUTER: IBM PC Compatible
OPERATING SYSTEM: MS-DOS

```

SOFTWARE: Word Processor (ASCII)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/426,436
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/705,557
FILING DATE:
APPLICATION NUMBER: US/08/057,531
FILING DATE:
APPLICATION NUMBER: 07/833,431
FILING DATE: February 7, 1992
ATTORNEY/AGENT INFORMATION:
NAME: Pasqualini, Patricia A.
REGISTRATION NUMBER: 34,894
REFERENCE/DOCKET NUMBER: 96700/238
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 697-5995
TELEFAX: (212) 286-0854 or 286-0082
TELEX: TWX 710-581-4766
INFORMATION FOR SEQ. ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 52297
TYPE: nucleotide
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE:
DESCRIPTION: phage genome sequence
HYPOTHETICAL: no
ANTI-SENSE: no
FRAGMENT TYPE: not applicable.
ORIGINAL SOURCE:
ORGANISM: mycobacteriophage L5
STRAIN: not applicable
INDIVIDUAL ISOLATE: L5
DEVELOPMENTAL STAGE: not applicable
HAPLOTYPE: not applicable
TISSUE TYPE: not applicable
CELL TYPE: not applicable
ORGANELLE: not applicable
IMMEDIATE SOURCE: mycobacteriophage L5 particles
POSITION IN GENOME: entire genome
FEATURE:
NAME/KEY:
LOCATION:
IDENTIFICATION METHOD:
OTHER INFORMATION:
PUBLICATION INFORMATION:
AUTHORS: Hatfull and Sarkis
TITLE: DNA Sequence, Structure and Gene
TITLE: Expression of Mycobacteriophage L5:
TITLE: A Phage System for Mycobacterial
TITLE: Genetics
JOURNAL: Molecular Microbiology
VOLUME: 7
PAGES: 395-405
DATE: 1993
US-09-426-436-1

```

Query Match 26.5%; Score 25.2; DB 4; Length 52297;
 Best Local Similarity 66.7%; Pred. No. 25;
 Matches 36; Conservative 0; Mismatches 18; Indels 0; Gaps 0;

QY 10 TCGCAGCTCTCTCAGAGAACAAAGACAGCCATTACCTGAGAGACTG 63
 ||||| ||| | | | | | | | | | | | | | | | | | | | |
 Db 47011 TCGCAGCTCTCTCGCCAGCGCAGAGACTGCGCGTACCGCGAAGAGGCTG 46958

Search completed: May 2, 2003, 12:49:30
 Job time : 56 secs

GenCore version 5.1.5
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 2, 2003, 12:25:30 ; Search time 63.2241 Seconds

(without alignments)
1775.493 Million cell updates/sec

Title: US-09-719-554-3_COPY_6956_7050

Perfect score: 95

Sequence: 1 ccctgctatcgcccaagctc.....accacacagcccaacctca 95

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 746064 seqs, 590810554 residues

Total number of hits satisfying chosen parameters: 1492128

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published_Applications_NA:*

1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/1/pubpna/PCF_NEW_PUB.seq:*
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/1/pubpna/PCFUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
10: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
11: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
12: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*
13: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
14: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	94	98.9	2930	10	US-09-902-535-1
2	94	98.9	2946	9	US-10-114-893-134
3	90.4	95.2	374	12	US-10-040-916-47
4	81.2	85.5	337	10	US-09-867-701-3363
5	71.2	74.9	758	7	US-08-979-847-112
6	30.8	32.4	414	10	US-09-864-761-25108
7	30.8	32.4	565	10	US-09-864-761-8377
8	29.4	30.9	222	10	US-09-974-300-8046
9	28.6	30.1	78056	9	US-10-109-551-1
10	28.6	30.1	1503841	9	US-09-946-807-1
11	28.6	30.1	1503841	10	US-09-795-666-1
12	28.6	30.1	1503841	10	US-09-795-666-1
13	27.8	29.3	3657	10	US-09-728-422-4
14	27.6	29.1	15070	10	US-09-764-877-3979
15	27.6	29.1	15071	10	US-09-764-877-3978
16	26.8	28.2	519	9	US-10-091-572-130
17	26.8	28.2	519	9	US-09-764-891-2099
18	26.8	28.2	8833	10	US-09-880-107-3272
19	26.2	27.6	7332	10	US-09-944-849-1

20	26.2	27.6	7673	10	US-09-815-925-1	Sequence 1, Appl1
21	26	27.4	1965	10	US-09-940-921B-3	Sequence 3, Appl1
22	26	27.4	2027	10	US-09-867-550-887	Sequence 887, App
23	26	27.4	2052	10	US-09-940-921B-1	Sequence 1, Appl1
24	26	27.4	2240	10	US-09-940-921B-5	Sequence 5, Appl1
25	26	27.4	2311	10	US-09-925-297-94	Sequence 94, Appl1
26	25.8	27.2	400	10	US-09-878-574-3631	Sequence 3631, Ap
27	25.8	27.2	3002	10	US-09-918-171A-1	Sequence 1, Appl1
28	25.8	27.2	5938	10	US-09-954-456-1607	Sequence 1607, Ap
29	25.8	27.2	5938	10	US-09-967-768A-225	Sequence 225, App
30	25.6	26.9	302	10	US-09-878-574-14590	Sequence 14590, A
31	25.4	26.7	329	10	US-09-960-352-8386	Sequence 8386, Ap
32	25.4	26.7	417	10	US-09-960-352-13655	Sequence 13655, A
33	25.4	26.7	450	9	US-09-918-995-25955	Sequence 25955, A
34	25.4	26.7	696	10	US-09-728-446-115	Sequence 115, App
35	25.4	26.7	859	10	US-09-925-300-683	Sequence 683, App
36	25.4	26.7	1939	9	US-09-925-299-89	Sequence 89, Appl1
37	25.4	26.7	1939	10	US-09-925-299-89	Sequence 89, Appl1
38	25.4	26.7	2272	9	US-10-174-590-345	Sequence 345, App
39	25.4	26.7	2272	9	US-10-176-758-345	Sequence 345, App
40	25.4	26.7	2272	9	US-10-175-737-345	Sequence 345, App
41	25.4	26.7	2272	9	US-10-173-706-345	Sequence 345, App
42	25.4	26.7	2272	9	US-10-175-738-345	Sequence 345, App
43	25.4	26.7	2272	9	US-10-175-752-345	Sequence 345, App
44	25.4	26.7	2272	9	US-10-176-482-345	Sequence 345, App
45	25.4	26.7	2272	9	US-10-176-757-345	Sequence 345, App

ALIGNMENTS

RESULT 1

US-09-902-535-1

Sequence 1, Application US/09902535

Patent No. US20020102530A1

GENERAL INFORMATION:

APPLICANT: Keith, Jr., James C.

APPLICANT: McCoy, John M.

APPLICANT: M. Sha

TITLE OF INVENTION: Methods and compositions for diagnosing and treating preclampsia and gestational trophoblast

TITLE OF INVENTION: disorders

FILE REFERENCE: GIN-6006B4

CURRENT APPLICATION NUMBER: US/09/902.535

CURRENT FILING DATE: 2001-07-09

PRIOR APPLICATION NUMBER: 60/216,657

PRIOR FILING DATE: 2000-07-06

NUMBER OF SEQ ID NOS: 5

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO: 1

LENGTH: 2930

TYPE: DNA

ORGANISM: Homo sapiens

FEATURE:

NAME/KEY: CDS

LOCATION: (930)...(2546)

US-09-902-535-1

Query Match

Best Local Similarity: 98.9%; Score 94; DB 10; Length 2930;

Matches 94; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 CTGCTTTTCCCAAGCTCTTCAGGAAACAAGAGCCATTACCTGGAGAAAC 61

DB 8 CTGCTTTTCCCAAGCTCTTCAGGAAACAAGAGCCATTACCTGGAGAAAC 67

QY 62 TGGCACTGATTATTCACCAAGCCCAACCTCA 95

DB 68 TGGCACTGATTATTCACCAAGCCCAACCTCA 101

RESULT 2

US-10-114-893-134

Sequence 134, Application US/10114893
Publication No. US20020193567A1
GENERAL INFORMATION:
APPLICANT: Jacobs, Kenneth
APPLICANT: McCoy, John M.
APPLICANT: Lavallee, Edward R.
APPLICANT: Collins-Racie, Lisa A.
APPLICANT: Evans, Cheryl
APPLICANT: Merberg, David
APPLICANT: Treacy, Maurice
APPLICANT: Bowman, Michael R.
APPLICANT: Spaulding, Vikki
APPLICANT: Carlin-Duckett, McKenough
APPLICANT: Kelleher, Kerry S.
APPLICANT: Genetics Institute, Inc.
TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES ENCODING THEM
FILE REFERENCE: GI 6000-10A
CURRENT APPLICATION NUMBER: US/10/114,893
EARLIER APPLICATION NUMBER: 09/413,232
EARLIER FILING DATE: 1999-10-06
NUMBER OF SEQ ID NOS: 321
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 134
LENGTH: 2946
TYPE: DNA
ORGANISM: Homo sapiens
US-10-114-893-134

Query Match 98.9%; Score 94; DB 9; Length 2946;
Best Local Similarity 100.0%; Pred. No. 3.5e-25;
Matches 94; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 CTGCTTATCGCAAGCTCTTCAGAGAAACAAGAGCCATTACCTGGAGAAGAC 61
DB 6 CTGCTTATCGCAAGCTCTTCAGAGAAACAAGAGCCATTACCTGGAGAAGAC 65
QY 62 TGGCAACTGATTTTACCCACAAGCCCAAACTCA 95
DB 66 TGGCAACTGATTTTACCCACAAGCCCAAACTCA 99

RESULT 3

US-10-040-916-47
Sequence 47, Application US/10040916
Patent No. US20020146769A1

GENERAL INFORMATION:

APPLICANT: Jacobs, Kenneth

McCoy, John

Lavallee, Edward

Racie, Lisa

Merberg, David

Treacy, Maurice

Evans, Cheryl

Spaulding, Vikki

TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES

ENCODING THEM

NUMBER OF SEQUENCES: 71

CORRESPONDENCE ADDRESS:

ADDRESSER: Genetics Institute, Inc.

STREET: 87 CambridgePark Drive

CITY: Cambridge

STATE: Massachusetts

COUNTRY: U.S.A.

ZIP: 02140

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA: US/10/040,916

APPLICATION NUMBER: US/10/040,916

FILING DATE: 07-Jan-2002

CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/887,029
FILING DATE: 07-FEB-1997
APPLICATION NUMBER: 08/686,878
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Brown, Scott A.
REGISTRATION NUMBER: 32,724
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 498-8224
TELEFAX: (617) 876-5851
INFORMATION FOR SEQ ID NO: 47:
SEQUENCE CHARACTERISTICS:
LENGTH: 374 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
SEQUENCE DESCRIPTION: SEQ ID NO: 47:
US-10-040-916-47

Query Match 95.2%; Score 90.4; DB 12; Length 374;
Best Local Similarity 96.8%; Pred. No. 4.1e-24;
Matches 91; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 CTGCTTATCGCAAGCTCTTCAGAGAAACAAGAGCCATTACCTGGAGAAGAC 61
DB 29 CTGCTTATCGCAAGCTCTTCAGAGAAACAAGAGCCATTACCTGGAAAAAANC 88
QY 62 TGGCAACTGATTTTACCCACAAGCCCAAACTCA 95
DB 89 TGGCAACTGATTTTACCCACAAGCCCAAACTCA 122

RESULT 4

US-09-867-701-3363/c
Sequence 3363, Application US/09867701
Patent No. US2002013237A1

GENERAL INFORMATION:

APPLICANT: Aglate, Paul A.

APPLICANT: Jones, Robert

APPLICANT: Harlocker, Susan L.

TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY

TITLE OF INVENTION: AND DIAGNOSIS OF OVARIAN CANCER

FILE REFERENCE: 210121.497

CURRENT APPLICATION NUMBER: US/09/867,701

NUMBER OF SEQ ID NOS: 10912

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 3363

LENGTH: 337

TYPE: DNA

ORGANISM: Homo sapien

FEATURE:

NAME/KEY: misc_feature

LOCATION: (1)...(337)

OTHER INFORMATION: n - A,T,C or G

US-09-867-701-3363

Query Match 85.5%; Score 81.2; DB 10; Length 337;
Best Local Similarity 90.5%; Pred. No. 1.1e-20;
Matches 86; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 1 CCGCTTATCGCAAGCTCTTCAGAGAAACAAGAGCCATTACCTGGAGAAGAC 60
DB 184 CCGCTTATCGCAAGCTCTTCAGAGAAACAAGAGCCATTACCTGGAGAAGAC 125
QY 61 CTGGCAACTGATTTTACCCACAAGCCCAAACTCA 95
DB 124 CTGGCAACTGATTTTACCCACAAGCCCAAACTCA 90

```
RESULT 5
US-08-979-847-112
; Sequence 112, Application US/08979847
; Publication No. US20030039664A1
; GENERAL INFORMATION:
; APPLICANT: PERRON, HERVE
; APPLICANT: BESEME, FREDERIC
; APPLICANT: BEDI, FREDERIC
; APPLICANT: PARANHOS-BACCALA, GLAUCIA
; APPLICANT: KOMORIAN-PRADEL, FLORENCE
; APPLICANT: JOLIVET-REYNAUD, COLETTE
; APPLICANT: MANDRAND, BERNARD
; APPLICANT: GARSON, JEREMY
; APPLICANT: TURE, PHILIP
; TITLE OF INVENTION: VIRAL MATERIAL AND NUCLEOTIDE FRAGMENTS
; TITLE OF INVENTION: ASSOCIATED WITH MULTIPLE SCLEROSIS, FOR DIAGNOSTIC, PROPHYLACT
; TITLE OF INVENTION: THERAPEUTIC PURPOSES
; NUMBER OF SEQUENCES: 210
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OLIVIER & BERRIDGE, PLC
; STREET: P.O. BOX 19928
; CITY: ALEXANDRIA
; STATE: VA
; COUNTRY: USA
; ZIP: 22320
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/979,847
; FILING DATE: 26-NOV-1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: BERRIDGE, WILLIAM P.
; REGISTRATION NUMBER: 30,024
; REFERENCE/DOCKET NUMBER: WPB 39046A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-836-6400
; TELEFAX: 703-836-2787
; INFORMATION FOR SEQ ID NO: 112:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 758 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; US-08-979-847-112

Query Match          74.9%; Score 71.2; DB 7; Length 758;
Best Local Similarity 90.6%; Pred. No. 7.6e-17;
Matches 8; Conservative 0; Mismatches 8; Indels 1; Gaps 1;

QY 1 CCGGCTTATGCGCAAGCTCTCTCAGAGAAAGAGGCGCATTCCTGGAGAGA 60
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 107 CCGGCTTATGCGCAAGCTCTCTCAGAGAAAGAGGCGCATTCCTGGAGAGA 166
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 61 CTGGCAACT-GATTTTACCAACAAGCCCAACTCA 95
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 167 CTGGCAACTGATTTTACCAACAAGCCCAACTCA 202
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 6
US-09-864-761-25108
; Sequence 25108, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
```

```
; TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
; FILE REFERENCE: Aeomica-x-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263,6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 25108
; LENGTH: 414
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC019191.2
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 4.4
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 2.6
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 2.7
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 3
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 2.2
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 2.4
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 2.9
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 3.4
; OTHER INFORMATION: SWISSPROT HIT: P21414, EVALU6 3.00e-16
; OTHER INFORMATION: NT HIT: Z78766.1, EVALU6 4.00e-62
; OTHER INFORMATION: EST_HUMAN HIT: BE142076.1, EVALU6 1.00e-59
US-09-864-761-25108

Query Match          32.4%; Score 30.8; DB 10; Length 414;
Best Local Similarity 61.0%; Pred. No. 0.077;
Matches 50; Conservative 0; Mismatches 32; Indels 0; Gaps 0;

QY 9 ATCCGCAAGCTCTTACGAGAGAACAAAGAGGCGCATTCCTGGAGAGAGCTGGCAAC 68
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 152 ATCATAGGCGCACTCTGGGGAACAAGAAATATGCACTATCCTGGAGAGAGCTGGCACT 211
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 69 TGATTTTACCAACAAGCCCAAA 90
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 212 TAGACTTCACCAATATGCTTAA 233
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

```

US-09-864-761-8377
: Sequence 8377, Application US/09864761
: Patent No. US20020048763A1
: GENERAL INFORMATION:
: APPLICANT: Penn, Sharon G.
: APPLICANT: Rank, David R.
: APPLICANT: Hanzel, David K.
: APPLICANT: Chen, Wensheng
: TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
: FILE REFERENCE: Aecm1ca-X-1
: CURRENT FILING DATE: 2001-05-23
: PRIOR APPLICATION NUMBER: US 60/180,312
: PRIOR FILING DATE: 2000-02-04
: PRIOR APPLICATION NUMBER: US 60/207,456
: PRIOR FILING DATE: 2000-05-26
: PRIOR APPLICATION NUMBER: US 09/632,366
: PRIOR FILING DATE: 2000-08-03
: PRIOR APPLICATION NUMBER: GB 24263.6
: PRIOR FILING DATE: 2000-10-04
: PRIOR APPLICATION NUMBER: US 60/236,359
: PRIOR FILING DATE: 2000-09-27
: PRIOR APPLICATION NUMBER: PCT/US01/00666
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00667
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00664
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00669
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00665
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00668
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00663
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00662
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00661
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00670
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: US 60/234,687
: PRIOR FILING DATE: 2000-09-21
: PRIOR APPLICATION NUMBER: US 09/608,408
: PRIOR FILING DATE: 2000-06-30
: PRIOR APPLICATION NUMBER: US 09/774,203
: PRIOR FILING DATE: 2001-01-29
: NUMBER OF SEQ ID NOS: 49117
: SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
: SEQ ID NO 8377
: LENGTH: 565
: TYPE: DNA
: ORGANISM: Homo sapiens
: FEATURE:
: OTHER INFORMATION: MAP TO AC019191.2
: OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 4.4
: OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 2.6
: OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 2.7
: OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 3
: OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 2.2
: OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 2.4
: OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 2.9
: OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 3.4
: US-09-864-761-8377
Query Match 32.4%; Score 30.8; DB 10; Length 565;
Best Local Similarity 61.0%; Pred. No. 0.085;
Matches 50; Conservative 0; Mismatches 32; Indels 0; Gaps 0;

```

```

OY      9  ATGGCCAGGCTCCTTCAGGAGAGAAAGACAGGCATTACCTCCGGGAGAACTGGCAAC 68
        ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||
Db      250 ATCTAAGGCCCACTCTGGGGGAAACAAAGAAATATG6CACTATCTCGAGACACTGGCACT 30
OY      69  TGATTTTACCCACAGGCCCAAA 90
        |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
Db      310 TAGACTTCACCCATATGCTCTAA 331

RESULT 8
US-09-974-300-8046
; Sequence 8046, Application US/09974300
; Patent No. US20020146721A1
; GENERAL INFORMATION:
; APPLICANT: Berka, Randy M.
; APPLICANT: Clausen, Ib Groth
; TITLE OF INVENTION: Methods For Monitoring Multiple Gene
; TITLE OF INVENTION: Expression
; FILE REFERENCE: 10085, 500-US
; CURRENT APPLICATION NUMBER: US/09/974,300
; CURRENT FILING DATE: 2001-10-05
; PRIOR APPLICATION NUMBER: 09/680,598
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: 60/279,526
; PRIOR FILING DATE: 2001-03-27
; NUMBER OF SEQ ID NOS: 8481
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8046
; LENGTH: 222
; TYPE: DNA
; ORGANISM: Bacillus clausii
; US-09-974-300-8046

```

Query Match	30.9%	Score 29.4	DB 10	Length 222
Best Local Similarity	60.8%	Pred. No. 0.21		
Matches 48	Conservative 0	Mismatches 31	Indels 0	Gaps 0
QY	5	CCTTTCGCCCAAGCTCCTTCAGAGACAAGAAACAGCGCCATTACCTCGAGAAAGACTGG	64	
Db	27	CCTCATCTCCAAAGCCCTTAATCGAGACGAAGACCATCTCGTTTCAGAGAACTAGATGA	86	
QY	65	CAACTGATTTTACCACAA	83	
Db	87	CAACCGATTTAAAGCGCAA	105	
RESULT 9				
US-10-109-551-1				
Sequence 1, Application US/10109551				
Publication No. US20020194635a1				
GENERAL INFORMATION:				
APPLICANT: DUNNE, PATRICK W.				
APPLICANT: PIEDRAHITA, JORGE				
TITLE OF INVENTION: TRANSGENIC ANIMALS RESISTANT TO TRANSMISSIBLE				
TITLE OF INVENTION: SPONGIFORM ENCEPHALOPATHIES				
FILE REFERENCE: TAMK:207US				
CURRENT APPLICATION NUMBER: US/10/109,551				
CURRENT FILING DATE: 2002-03-28				
PRIOR APPLICATION NUMBER: 60/280,549				
PRIOR FILING DATE: 2001-03-30				
NUMBER OF SEQ ID NOS: 10				
SOFTWARE: PatentIn Ver. 2.1				
SEQ ID NO 1				
LENGTH: 78056				
TYPE: DNA				
ORGANISM: Bos taurus				
US-10-109-551-1				
Query Match	30.1%	Score 28.6	DB 9	Length 78056
Best Local Similarity	61.3%	Pred. No. 2.6		
Matches 46	Conservative 0	Mismatches 29	Indels 0	Gaps 0
Y	16	AGCTTCCTTCAGGAGACAAAGACAGCGCCATTACCTCGAGAGAGACTGCGCAACTGATT	75	

Db 29105 AGCAAAATCTGCTTACATGAGTGAATATCTGAGACGACCTGGAAATTGATTT 29164

QY 76 ACCCAACAAGCCCA 90

Db 29165 TACTCATATGCCAA 29179

RESULT 10

US-09-946-807-1/c
; Sequence 1, Application US/09946807
; Patent No. US20020165144A1
; GENERAL INFORMATION:
; APPLICANT: Stefansson, Hreinn
; APPLICANT: Steinthorsdottir, Valgerdur
; APPLICANT: Gulcher, Jeffrey R.
; TITLE OF INVENTION: HUMAN SCHIZOPHRENIA GENE
; FILE REFERENCE: 2345.2004-001
; CURRENT APPLICATION NUMBER: US/09/946,807
; CURRENT FILING DATE: 2001-09-05
; PRIOR APPLICATION NUMBER: US/09/795,668
; PRIOR FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: US 09/515,716
; PRIOR FILING DATE: 2000-02-28
; NUMBER OF SEQ ID NOS: 1531
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 1503841
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(1531)
; OTHER INFORMATION: y=t/u or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(1531)
; OTHER INFORMATION: r=g or a
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(1531)
; OTHER INFORMATION: m=a or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(1531)
; OTHER INFORMATION: k=g or t/u
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(1531)
; OTHER INFORMATION: s=g or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(1531)
; OTHER INFORMATION: w=a or t/u
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(1531)
; OTHER INFORMATION: d=a or g or t/u
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(1531)
; OTHER INFORMATION: h=a or c or t/u
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(1531)
; OTHER INFORMATION: v=a or g or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(1531)

; OTHER INFORMATION: n=a or g or c or t/u
US-09-946-807-1

Query Match 30.1%; Score 28.6; DB 9; Length 1503841;
Best Local Similarity 64.2%; Pred. No. 6.4;
Matches 43; Conservative 0; Mismatches 24; Indels 0; Gaps 0;

QY 13 CCAAGCTCCCTTCAGGAGCAACAGACGACCATACCTGGAGAGACTGGCAACTGAT 72

Db 784413 CGAAATCTTCAGTAGAGAAACAAATGTAATTCCTGGAGAGATGCTCCCTCTGT 784354

QY 73 TTACCC 79

Db 784353 ATTGCC 784347

RESULT 11

US-09-795-668-1/c
; Sequence 1, Application US/09795668
; Patent No. US20020045577A1
; GENERAL INFORMATION:
; APPLICANT: Stefansson, Hreinn
; APPLICANT: Steinthorsdottir, Valgerdur
; APPLICANT: Gulcher, Jeffrey R.
; TITLE OF INVENTION: HUMAN SCHIZOPHRENIA GENE
; FILE REFERENCE: 2345.2004-001
; CURRENT APPLICATION NUMBER: US/09/795,668
; CURRENT FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: US 09/515,716
; PRIOR FILING DATE: 2000-02-28
; NUMBER OF SEQ ID NOS: 1531
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 1503841
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(1531)
; OTHER INFORMATION: r=g or a
; NAME/KEY: misc_feature
; LOCATION: (1)...(1531)
; OTHER INFORMATION: y=t/u or c
; NAME/KEY: misc_feature
; LOCATION: (1)...(1531)
; OTHER INFORMATION: m=a or c
; NAME/KEY: misc_feature
; LOCATION: (1)...(1531)
; OTHER INFORMATION: k=g or t/u
; NAME/KEY: misc_feature
; LOCATION: (1)...(1531)
; OTHER INFORMATION: s=g or c
; NAME/KEY: misc_feature
; LOCATION: (1)...(1531)
; OTHER INFORMATION: w=a or t/u
; NAME/KEY: misc_feature
; LOCATION: (1)...(1531)
; OTHER INFORMATION: d=a or g or t/u
; NAME/KEY: misc_feature
; LOCATION: (1)...(1531)
; OTHER INFORMATION: h=a or c or t/u
; NAME/KEY: misc_feature
; LOCATION: (1)...(1531)
; OTHER INFORMATION: v=a or g or c
; NAME/KEY: misc_feature
; LOCATION: (1)...(1531)
; OTHER INFORMATION: n=a or g or c or t/u
US-09-795-668-1

Query Match 30.1%; Score 28.6; DB 10; Length 1503841;

Best Local Similarity 64.2%; Pred. No. 6.4;
Matches 43; Conservative 0; Mismatches 24; Indels 0; Gaps 0;

QY 13 CCAAGCTCCTTCAGAGAACAAAGACAGCCATTACCTGGAGAGACTGGCAACTGAT 72

DB 784413 CGAAATTCCTTCAGTAGTAGAACAATGTATTCTCGGAGAGTCTCCCTCTGT 784354

QY 73 TTATCCC 79

DB 784353 ATTTCCC 784347

RESULT 12

US-09-795-686-1/c
Sequence 1, Application US/09795686
Patent No. US20020094954A1

GENERAL INFORMATION:

APPLICANT: Steiansson, Hrelnn

APPLICANT: Steinthorsdottir, Valgerdur

APPLICANT: Gulcher, Jeffrey R.

TITLE OF INVENTION: HUMAN SCHIZOPHRENIA GENE

FILE REFERENCE: 2345-2005-001

CURRENT APPLICATION NUMBER: US/09/795,686

CURRENT FILING DATE: 2001-02-28

PRIOR APPLICATION NUMBER: US 09/515,715

PRIOR FILING DATE: 2000-02-28

NUMBER OF SEQ ID NOS: 1531

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 1

LENGTH: 1503841

TYPE: DNA

ORGANISM: Homo sapiens

FEATURE:

NAME/KEY: misc_feature

LOCATION: (1)...(1531)

OTHER INFORMATION: r-g or a

NAME/KEY: misc_feature

LOCATION: (1)...(1531)

OTHER INFORMATION: y-t/u or c

NAME/KEY: misc_feature

LOCATION: (1)...(1531)

OTHER INFORMATION: m-a or c

NAME/KEY: misc_feature

LOCATION: (1)...(1531)

OTHER INFORMATION: k-g or t/u

NAME/KEY: misc_feature

LOCATION: (1)...(1531)

OTHER INFORMATION: s-g or c

NAME/KEY: misc_feature

LOCATION: (1)...(1531)

OTHER INFORMATION: w-a or t/u

NAME/KEY: misc_feature

LOCATION: (1)...(1531)

OTHER INFORMATION: b-g or c or t/u

NAME/KEY: misc_feature

LOCATION: (1)...(1531)

OTHER INFORMATION: d-a or g or t/u

NAME/KEY: misc_feature

LOCATION: (1)...(1531)

OTHER INFORMATION: h-a or c or t/u

NAME/KEY: misc_feature

LOCATION: (1)...(1531)

OTHER INFORMATION: v-a or g or c

NAME/KEY: misc_feature

LOCATION: (1)...(1531)

OTHER INFORMATION: n-a or g or c or t/u

US-09-795-686-1

Query Match

Best Local Similarity 30.1%; Score 28.6; DB 10; Length 1503841;

Matches 43; Conservative 0; Mismatches 24; Indels 0; Gaps 0;

QY 13 CCAAGCTCCTTCAGAGAACAAAGACAGCCATTACCTGGAGAGACTGGCAACTGAT 72

DB 784413 CGAAATTCCTTCAGTAGTAGAACAATGTATTCTCGGAGAGTCTCCCTCTGT 784354

QY 73 TTATCCC 79

DB 784353 ATTTCCC 784347

RESULT 13

US-09-728-422-4
Sequence 4, Application US/09728422
Patent No. US20020128187A1

GENERAL INFORMATION:

APPLICANT: Tang, Y. Tom

APPLICANT: Zhou, Ping

APPLICANT: Goodrich, Ryle

APPLICANT: Liu, Chenghua

APPLICANT: Asundi, Vinod

APPLICANT: Ren, Feiyun

APPLICANT: Zhao, Qing A.

APPLICANT: Yang, Yonchong

APPLICANT: Wehtman, Tom

APPLICANT: Drmanac, Radoje T.

TITLE OF INVENTION: No. US20020128187A1el Nucleic Acids and

FILE REFERENCE: 787CIP2F

CURRENT APPLICATION NUMBER: US/09/728,422

CURRENT FILING DATE: 2000-11-30

PRIOR APPLICATION NUMBER: 09/560,875

PRIOR FILING DATE: 2000-04-27

PRIOR APPLICATION NUMBER: 09/496,914

PRIOR FILING DATE: 2000-02-03

NUMBER OF SEQ ID NOS: 10

SOFTWARE: pl-Fl-Genes Version 2.0

SEQ ID NO 4

LENGTH: 3657

TYPE: DNA

ORGANISM: Homo sapiens

FEATURE:

NAME/KEY: CDS

LOCATION: (645)..(3254)

US-09-728-422-4

Query Match

Best Local Similarity 29.3%; Score 27.8; DB 10; Length 3657;

Matches 47; Conservative 0; Mismatches 32; Indels 0; Gaps 0;

QY 9 ATGCCAGCTCCTTCAGAGAACAAAGACAGCCATTACCTGGAGAGACTGGCAAC 68

DB 744 ATCTTAAGATCAATGACATACAGAGTAGATGGCCTTCAGATCTCTGAGAC 803

QY 69 TGATTTTACCACAGGCC 87

DB 804 TTGTTAGCAACAAGTCC 822

RESULT 14

US-09-764-877-3979/c
Sequence 3979, Application US/09764877
Patent No. US20020147140A1

GENERAL INFORMATION:

APPLICANT: Rosen et al.

TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies

FILE REFERENCE: PC005

CURRENT APPLICATION NUMBER: US/09/764,877

CURRENT FILING DATE: 2001-01-17

Prior application data removed - refer to PALM or file wrapper

NUMBER OF SEQ ID NOS: 4031

SOFTWARE: PatentIn Ver. 2.0

SEQ ID NO 3979

LENGTH: 15070

TYPE: DNA

ORGANISM: Homo sapiens

US-09-764-877-3979

Query Match 29.1%; Score 27.6; DB 10; Length 15070;
Best Local Similarity 63.6%; Pred. No. 3.7;
Matches 42; Conservative 0; Mismatches 24; Indels 0; Gaps 0;

QY 25 AGGAGAACAAAGACAGCCATTACCTGAGAGAGACTGGCACTGATTTTACCCACAAG 84

Db 6635 AGAAAAAGAAAAAGAAAAAGATATGAAACAGACAGACACATCATGTGGCCCAAAAT 6576

QY 85 CCCAAA 90

Db 6575 GCCTAA 6570

RESULT 15

US-09-764-877-3978/C
; Sequence 3978, Application US/09764877

; Patent No. US20020147140A1

; GENERAL INFORMATION:

; APPLICANT: Rosen et al.

; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies

; FILE REFERENCE: PC005

; CURRENT APPLICATION NUMBER: US/09/764,877

; CURRENT FILING DATE: 2001-01-17

; Prior application data removed - refer to PALM or file wrapper

; NUMBER OF SEQ ID NOS: 4031

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 3978

; LENGTH: 15071

; TYPE: DNA

; ORGANISM: Homo sapiens

US-09-764-877-3978

Query Match 29.1%; Score 27.6; DB 10; Length 15071;
Best Local Similarity 63.6%; Pred. No. 3.7;
Matches 42; Conservative 0; Mismatches 24; Indels 0; Gaps 0;

QY 25 AGGAGAACAAAGACAGCCATTACCTGAGAGAGACTGGCACTGATTTTACCCACAAG 84

Db 6635 AGAAAAAGAAAAAGAAAAAGATATGAAACAGACAGACACATCATGTGGCCCAAAAT 6576

QY 85 CCCAAA 90

Db 6575 GCCTAA 6570

Search completed: May 2, 2003, 15:02:30
Job time : 1219.22 secs

GenCore version 5.1.5
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 2, 2003, 12:54:45 ; Search time 689 Seconds

(without alignments)
8490.078 Million cell updates/sec

Title: US-09-719-554-3_COPY-7500-7700

Perfect score: 201
Sequence: 1 caactagtcgacacatca.....atttattcaccacactgc 201

Scoring table: IDENTITY_NDC
Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

GenEmbl: 1: gb_ba: 2: gb_htg: 3: gb_in: 4: gb_om: 5: gb_ov: 6: gb_pat: 7: gb_ph: 8: gb_pl: 9: gb_pr: 10: gb_ro: 11: gb_sts: 12: gb_sy: 13: gb_un: 14: gb_vi: 15: em_ba: 16: em_fun: 17: em_hum: 18: em_in: 19: em_mu: 20: em_om: 21: em_or: 22: em_ov: 23: em_ph: 24: em_pl: 25: em_ro: 26: em_sts: 27: em_un: 28: em_vl: 29: em_vt: 30: em_htg_hum: 31: em_htg_inv: 32: em_htg_other: 33: em_htg_mus: 34: em_htg_pln: 35: em_htg_rod: 36: em_htg_mam: 37: em_htg_vrt: 38: em_higo_hum: 39: em_higo_mus: 40: em_higo_other: 41: em_higo_other:

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	201	100.0	2055	6	AX007999 Sequence
2	201	100.0	2599	6	AX007978 Sequence
3	201	100.0	2930	6	AX355872 Sequence
4	201	100.0	2930	9	AF208161 Homo sapi
5	201	100.0	2946	6	AR177269 Sequence
6	201	100.0	10499	6	AX007980 Sequence
7	201	100.0	56093	6	AX329572 Sequence
8	201	100.0	56093	9	HSAC000064
9	201	100.0	149194	9	AC007566
10	199.4	99.2	7582	6	AX000966 Sequence
11	199.4	99.2	7582	6	AX027480
12	197.8	98.4	2006	6	AX000959 Sequence
13	197.8	98.4	2006	6	AX027473
14	197.8	98.4	2006	9	AF072503
15	185	92.0	2477	14	AF123882
16	185	92.0	158033	9	AC018926
17	185	92.0	187321	2	AC092510
18	184	91.5	155021	9	HS774G10
19	180.4	89.8	159851	9	AC036114
20	180.4	89.8	16256	6	AC018922
21	180.4	89.8	195344	2	AC025248
22	180.2	89.7	2575	6	AX000969 Sequence
23	180.2	89.7	46575	2	AC080036
24	180.2	89.7	114621	9	AC005187
25	180.2	89.7	170754	2	AC012403
26	180.2	89.7	177866	2	AL354698 Homo sapi
27	180.2	89.7	180954	9	AL160281 Human DNA
28	179	89.1	3372	6	AX002478 Sequence
29	179	89.1	3372	9	AF072502
30	179	89.1	3372	2	AC009727
31	178.8	89.0	188607	2	AC008121
32	178.8	89.0	200926	2	AC008121
33	178.6	88.9	1216	6	AX007984
34	178.6	88.9	2784	6	AX007981
35	178.6	88.9	103056	9	AC068447
36	178.6	88.9	135734	2	AC002346
37	178.6	88.9	161812	9	HS197017
38	178.6	88.9	177476	9	AC015983
39	178.6	88.9	195032	9	AL138706 Human DNA
40	178.6	88.9	205588	9	AC068400
41	177.2	88.2	88823	9	AL713965 Human DNA
42	177	88.1	153444	2	AC040948
43	177	88.1	205035	9	CNS00005
44	172.6	85.9	73070	9	AL590143 Human chr
45	172.6	85.9	160883	2	AL158814 Homo sapi

ALIGNMENTS

RESULT 1	AX007999	2055 bp	DNA	linear	PAT 06-SEP-2000
LOCUS	AX007999				
DEFINITION	Sequence 22 from Patent WO967395.				
ACCESSION	AX007999				
VERSION	AX007999.1	GI:9995696			
KEYWORDS					
SOURCE	human.				
ORGANISM	Homo sapiens				

REFERENCE 1 (bases 1 to 2055)
Perin,J.P., Rieger,F. and Alliel,P.M.
Nucleic sequence and deduced protein sequence family with human
endogenous retroviral motifs, and their uses

Pred. No. is the number of results predicted by chance to have a

JOURNAL

Patent: WO 9967395-A 22 29-DEC-1999;
INST NAT SANTE RECH MED (FR); PERIN JEAN PIERRE (FR); RIEGER
FRANCOIS (FR); ALLIEL PATRICK M (FR)

FEATURES

Location/Qualifiers
1..2055

CDS

/organism="Homo sapiens"
/db_xref="taxon:9606"
<1..87
/note="unnamed protein product"
/codon_start=1
/protein_id="CAC07392.1"
/db_xref="GI:9995697"
/translation="PRTANLVADITSLAKYQVLTQSTYP"

BASE COUNT

576 a 574 c 376 g 529 t

ORIGIN

Query Match 100.0%; Score 201; DB 6; Length 2055;
Best Local Similarity 100.0%; Pred. No. 9.3e-52;
Matches 201; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY

1 CAACTAGTTGACAGATCCTCCTTAGCCAAATATCAAGTTCTTAAACATTACA 60

DB

12 CAACTAGTTGACAGATCCTCCTTAGCCAAATATCAAGTTCTTAAACATTACA 71

QY

61 AGAACCTATCCCTGAGAGAGGAGAAAGACATATTCACCTGTGACATGTTAGT 120

DB

72 AGAACCTATCCCTGAGAGAGGAGAAAGACATATTCACCTGTGACATGTTAGT 131

QY

121 CAAGTCCTTCCCTTAATTCCTCCATGCCATGATCTCTGGAGAGACCTACCCAGT 180

DB

132 CAAGTCCTTCCCTTAATTCCTCCATGCCATGATCTCTGGAGAGACCTACCCAGT 191

QY

181 CATTATATCTACCCCAACTGC 201

DB

192 CATTATATCTACCCCAACTGC 212

RESULT 2

AX007978 2599 bp DNA linear PAT 06-SEP-2000
DEFINITION Sequence 1 from Patent WO9967395.
ACCESSION AX007978
VERSION AX007978.1 GI:9995675
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens

REFERENCE

1 (bases 1 to 2599)
Perlin,J.P., Rieger,F. and Alliel,P.M.
Nucleic sequence and deduced protein sequence family with human
endogenous retroviral motifs, and their uses
Patent: WO 9967395-A 1 29-DEC-1999;
INST NAT SANTE RECH MED (FR); PERIN JEAN PIERRE (FR); RIEGER
FRANCOIS (FR); ALLIEL PATRICK M (FR)

JOURNAL

Location/Qualifiers
1..2599
/organism="Homo sapiens"
/db_xref="taxon:9606"

BASE COUNT

744 a 718 c 495 g 642 t

ORIGIN

Query Match 100.0%; Score 201; DB 6; Length 2599;
Best Local Similarity 100.0%; Pred. No. 9.5e-52;
Matches 201; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY

1 CAACTAGTTGACAGATCCTCCTTAGCCAAATATCAAGTTCTTAAACATTACA 60

DB

550 CAACTAGTTGACAGATCCTCCTTAGCCAAATATCAAGTTCTTAAACATTACA 609

QY

61 AGAACCTATCCCTGAGAGAGGAGAAAGACATATTCACCTGTGACATGTTAGT 120

DB

610 AGAACCTATCCCTGAGAGAGGAGAAAGACATATTCACCTGTGACATGTTAGT 669

QY 121 CAAGTCCTTCCCTTAATTCCTCCATGCCATGATCTCTGGAGAGACCTACCCAGT 180
DB 670 CAAGTCCTTCCCTTAATTCCTCCATGCCATGATCTCTGGAGAGACCTACCCAGT 729
QY 181 CATTATATCTACCCCAACTGC 201
DB 730 CATTATATCTACCCCAACTGC 750

RESULT 3

AX355872 2930 bp DNA linear PAT 06-FEB-2002
LOCUS AX355872
DEFINITION Sequence 1 from Patent WO0204678.
ACCESSION AX355872
VERSION AX355872.1 GI:18620523
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens

REFERENCE

1 Keith,J.C., McCoy,J.M. and M.S.
Methods and compositions for diagnosing and treating preeclampsia
and gestational trophoblast disorders
Patent: WO 0204678-A 1 17-JAN-2002;
GENETICS INSTITUTE, INC. (US)

JOURNAL

Location/Qualifiers
1..2930
/organism="Homo sapiens"
/db_xref="taxon:9606"

FEATURES

Source human.
Organism Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.

CDS

/note="unnamed protein product"
/codon_start=1
/protein_id="CAD23000.1"
/db_xref="GI:18620524"

/translation="MALPYHIFLFTVLPSFTLAPPCRCMTSSPYQEFILRMQRP
GNIDAPYSRLSKGPTPTFAHMPNRCYHSATLGMHANTHYTKMLNPSCPGILGY
IVCMYFTQGMDSGGVODQAREKHEKVIQSLRVHGTSSPYGLDLSKHEHETLT
HRLVLSLENTLTGLHEVSAONPTNMCILPLNRPYVSIIPVEDMNNSTFENITTSV
LVGPLVSNLEITHNTSLNLCVRSNTNTTNSCINFWPPPOIVLPSIFPVCJSA
YKLGSSSESKFSLFVPEPTITTEODLYNVIISPRKRVPLFVYGAIVLALG
TGIGITFTTGFYKISDELNGDMERVAQSLVLDQLSLAVALVQNRALDLTAE
RGCTGLFLEECYVYNOGIVTEKVEIKRDRIRORAEIRNTGFMGLSGMPPILP
FLGPLAAILILLFPCIFNLVNFVSSRIEAVKIQMEBKMSKRIYRRPDRASP
RSDVNDIKTPPEEISAAQPLLRPSAGSS"

BASE COUNT

842 a 800 c 571 g 717 t

ORIGIN

Query Match 100.0%; Score 201; DB 6; Length 2930;
Best Local Similarity 100.0%; Pred. No. 9.6e-52;
Matches 201; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY

1 CAACTAGTTGACAGATCCTCCTTAGCCAAATATCAAGTTCTTAAACATTACA 60

DB

551 CAACTAGTTGACAGATCCTCCTTAGCCAAATATCAAGTTCTTAAACATTACA 610

QY

61 AGAACCTATCCCTGAGAGAGGAGAAAGACATATTCACCTGTGACATGTTAGT 120

DB

611 AGAACCTATCCCTGAGAGAGGAGAAAGACATATTCACCTGTGACATGTTAGT 670

QY

121 CAAGTCCTTCCCTTAATTCCTCCATGCCATGATCTCTGGAGAGACCTACCCAGT 180

DB

671 CAAGTCCTTCCCTTAATTCCTCCATGCCATGATCTCTGGAGAGACCTACCCAGT 730

QY

181 CATTATATCTACCCCAACTGC 201

DB

731 CATTATATCTACCCCAACTGC 751

RESULT 4

AF208161 2930 bp mRNA linear PRI 22-FEB-2000
LOCUS AF208161

BASE COUNT 3048 a 2676 c 2280 g 2495 t
 ORIGIN

Query Match 100.0%; Score 201; DB 6; Length 10499;
 Best Local Similarity 100.0%; Pred. No. 1.1e-51;
 Matches 201; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CAACCTAGTGGACAGATCCTCCTAGCAATATACAACTGCTTAAACATTACA 60
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 DB 7500 CAACCTAGTGGACAGATCCTCCTAGCAATATACAACTGCTTAAACATTACA 7559
 QY 61 AGGAACCTATCCCTGAGAGAGGAGAAAGAACTATTCACCTTGTGACATGATTAGT 120
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 DB 7560 AGGAACCTATCCCTGAGAGAGGAGAAAGAACTATTCACCTTGTGACATGATTAGT 7619
 QY 121 CAAGTCCCTTCCCTTAATTCCTCCATCCCTAGATACATCTCTGGAGAGACCTACCCAGT 180
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 DB 7620 CAAGTCCCTTCCCTTAATTCCTCCATCCCTAGATACATCTCTGGAGAGACCTACCCAGT 7679
 QY 181 CATTATCTACCCCAACTGC 201
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 DB 7680 CATTATCTACCCCAACTGC 7700

RESULT 7
 AX329572 56093 bp DNA linear PAT 09-JAN-2002
 LOCUS AX329572
 DEFINITION Sequence 81 from Patent WO0194629.
 ACCESSION AX329572
 VERSION AX329572.1 GI:18102550
 KEYWORDS
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1
 Young, P. E., Augustus, M., Carter, K. C., Ebner, R., Endress, G.,
 Horrigan, S., Soppet, D. R. and Weaver, Z.
 TITLE Cancer gene determination and therapeutic screening using signature
 gene sets
 JOURNAL Patent: WO 0194629-A 81 13-DEC-2001;
 Avalon Pharmaceuticals (US)
 FEATURES
 source 1. 56093
 Location/Qualifiers
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 BASE COUNT 16164 a 12346 c 10702 g 16881 t
 ORIGIN

Query Match 100.0%; Score 201; DB 6; Length 56093;
 Best Local Similarity 100.0%; Pred. No. 1.2e-51;
 Matches 201; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CAACCTAGTGGACAGATCCTCCTAGCAATATACAACTGCTTAAACATTACA 60
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 DB 35500 CAACCTAGTGGACAGATCCTCCTAGCAATATACAACTGCTTAAACATTACA 35559
 QY 61 AGGAACCTATCCCTGAGAGAGGAGAAAGAACTATTCACCTTGTGACATGATTAGT 120
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 DB 35560 AGGAACCTATCCCTGAGAGAGGAGAAAGAACTATTCACCTTGTGACATGATTAGT 35619
 QY 121 CAAGTCCCTTCCCTTAATTCCTCCATCCCTAGATACATCTCTGGAGAGACCTACCCAGT 180
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 DB 35620 CAAGTCCCTTCCCTTAATTCCTCCATCCCTAGATACATCTCTGGAGAGACCTACCCAGT 35679
 QY 181 CATTATCTACCCCAACTGC 201
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 DB 35680 CATTATCTACCCCAACTGC 35700

RESULT 8

HSAC000064
 LOCUS HSAC000064 56093 bp DNA linear PRI 13-NOV-1996
 DEFINITION Human BAC clone RG083M05 from 7q21-7q22, complete sequence.
 AC000064
 VERSION AC000064.1 GI:1669369
 KEYWORDS HTG.
 SOURCE Homo sapiens.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 56093)
 PAULEY, A.
 TITLE The sequence of H. sapiens BAC clone RG083M05
 JOURNAL Unpublished (1996)
 REFERENCE 2 (bases 1 to 56093)
 WATERSTON, R.
 TITLE Direct Submission
 JOURNAL Submitted (13-NOV-1996)
 COMMENT Genome Sequencing Center
 Department of Genetics, Washington University
 St. Louis, MO 63108, USA
 e-mail: saplens@watson.wustl.edu

NOTICE: This sequence may not represent the entire insert of this clone. It may be shorter because we only sequence overlapping sections once, or longer because we provide a small overlap between neighboring submissions.

This sequence was finished as follows unless otherwise noted:
 all regions were double stranded or sequenced with an alternate chemistry; an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

SOURCE INFORMATION:
 This clone is from the first release of the human BAC library. The library contains cloned DNA from a human male fibroblast cell line 978SK. For references see: Shizuya et al., Proc. Natl. Acad. Sci. 89:8794-8797 (1992); Kim et al., Genomics 34:213-218 (1996).
 VECTOR: pBELO
 Selection: chloramphenicol

NEIGHBORING SEQUENCE INFORMATION:
 The orientation of this clone is unknown. Actual start of this clone is at base position 1 of H_RG083M05; actual end is at 56093 of H_RG083M05

This clone contains STS SWS1725.

FEATURES
 source 1. 56093
 Location/Qualifiers
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="7"
 /map="7q21-7q22"
 /clone="H_RG083M05"
 /clone_1kb="CTP-978SK-B"
 complement(838..1131)
 /rpl_family="ALU"
 /rpl_family="ALU"
 /gene="WUSC:H_RG083M05.1"
 /gene="WUSC:H_RG083M05.1"
 /gene="WUSC:H_RG083M05.1"
 /gene="WUSC:H_RG083M05.1"
 /note="ATPase; strong similarity to peroxisome
 biosynthesis protein PAB1 (PID:G1172019); coded for by
 human cDNA C04279 (NID:g1467530)"
 /codon_start=1
 /protein_id="BAB46346.1"
 /db_xref="GI:1669371"
 /translation="KRLNIIQKTLLEAFSEAVMOPSVLLDLDLGLPAVPEHEH
 SPDAQRCEIILCNVINKKIDCDINKFTDLDLGHVAKETGFAVRFYTVLVDAIRHSRL

```
SNOS1STEKCEVLTITDPOKALRGCLPASLSYVLNHRPRDLGMNKTIQGLHEVRILMD  
TIOUAKCYLKEKKETPELPANLPFRTGIGLLYGPGTGKTLLAGVTARESRNFISVQ  
KGPELLSYKIASEEQAVRDIFIRAOAAAPCILEPDEDEFSLAPRGHNTGYTVENNO  
LIOTDGEGVGOGVVLAATSRPDIDPALHRLDKCVYCPCPPDOVTTJESLKRO  
MLOHSFIWLEIIVNLSDSLEPADVDJOHVASVTSOFTCADIKALLYNAQLALKMG  
MLDSKFELIPDESCKFNMYRLPYGSSESELGNSSDLSGCCIASPSNMODLGPVP  
EKDDQFSOPPVLRTRSGCGCELTOEORDOLRADIIITIKNTRYSCGDESMNQPGPI  
KRRIALSOSHMTLAGHTRPBISEDMMKNFAEL"
```

```
repeat_region complement(49.8, .5130)  
    /rpt_family="Alu"  
repeat_region complement(6581, ..7133)  
    /rpt_family="t1"  
repeat_region complement(7767, .8037)  
    /rpt_family="Alu"  
repeat_region complement(8186, ..8472)  
misc_feature      /rpt_family="Alu"  
                    8473..8625  
misc_feature       /gene="WUGSC:H.RG083M05.1"  
note="match to human 3' EST H75762 (NID:g1049794), bases  
287-444"  
                    8841..9161  
misc_feature       /gene="WUGSC:H.RG083M05.1"  
note="match to human 5' EST H75921 (NID:g1050050), bases  
21-348"  
misc_feature        9481..9547  
                     /gene="WUGSC:H.RG083M05.1"  
note="match to human 5' EST N22627 (NID:g1130501), bases  
276-743"  
repeat_region     complement(12612, ..12907)  
misc_feature      /rpt_family="Alu"  
                    13670..13793  
repeat_region     /gene="WUGSC:H.RG083M05.1"  
note="match to human 5' EST H41382 (NID:917434), bases  
143-266"  
                  13794..13877  
repeat_region     /rpt_family="Alu"  
                    13878..13906  
misc_feature       /gene="WUGSC:H.RG083M05.1"  
note="match to human 5' EST H41382 (NID:g917434), bases  
30-58"  
repeat_region     13907..14104  
                     /rpt_family="Alu"  
repeat_region     complement(14110, .14137)  
                     /rpt_family="t1"  
repeat_region     complement(15618, .15907)  
repeat_region     /rpt_family="Alu"  
                    17227..17522  
misc_feature       /rpt_family="Alu"  
                    18667..19235  
repeat_region     note="match to human fetal brain 5' EST D61494  
                    (NID:g970409), bases 1-255, and to human 3' EST R07476  
                    (NID:g759399)"  
                    19550..19670  
repeat_region     /rpt_family="Alu"  
                    21507..37303  
repeat_region     note="similarity to various SS-RNA virus polypeptides;  
                    pseudogene; region of matches and close matches to  
                    multiple human ESTs, see R68740 (NID:g842257)"  
                    exon          37316..37489  
  
repeat_region     /note="Grail prediction, score = 80"  
                    evidence=not_experimental  
repeat_region     complement(38938, .39224)  
misc_feature      /rpt_family="Alu"  
                    39225..39707  
misc_feature       /note="match to multiple human ESTs, see N30113  
                    (NID:g1148633)"  
                    39800..40085  
repeat_region     /rpt_family="Alu"  
                    repeat_region complement(40247, .40538)  
                        /rpt_family="Alu"  
                        complement(40632, .40924)  
repeat_region     /rpt_family="Alu"  
                    complement(42283, .42891)
```

	repeat_region	/rpt_family="ALU"	complement(45474 .45613)
	misc_feature	/rpt_family="ALU"	complement(45614 .45737)
	misc_feature	/note="match to human 3' EST H48898 (NID:g988738), bases 129-333"	
	misc_feature	complement(46107 .47026)	
		/note="match to multiple human ESTs, see NB1064 (NID:g1243765), H48897 (NID:g988737), and M78831 (NID:g273146)"	
	repeat_region	complement(47027 .47318)	
	misc_feature	/rpt_family="ALU"	complement(47365 .47782)
	misc_feature	/note="match to multiple human ESTs, see W37495 (NID:g1319089)"	
	misc_feature	47898 .48115	
		/note="match to human 5' EST H62306 (NID:g1015138), bases 93-368"	
	repeat_region	complement(48116 .48405)	
	misc_feature	/rpt_family="ALU"	complement(48406 .48584)
		/note="match to human 3' EST N29952 (NID:g1148472), bases 290-455, and 5' EST R12730 (NID:g765806)"	
	repeat_region	complement(48787 .49405)	
	misc_feature	/rpt_family="ALU"	complement(49406 .49534)
		/note="match to human 3' EST R65794 (NID:g838432), bases 309-440"	
	repeat_region	complement(49638 .49672)	
	misc_feature	/rpt_family="ALU"	complement(49674 .49890)
		/note="match to human 3' EST N29938 (NID:g1148458), sequences are from opposite ends of the same clone"	
	gene	complement(49698 .51806)	
	CDS	/gene="WUSC:H.RG083M05.2"	
		complement(join(49698 .49888,51575 .51806))	
		/note="coded for by human cDNAs W37389 (NID:g1319205), R65891 (NID:g838529), R65794 (NID:g838432) and R65794 (NID:g838432)"	
		/codon_start=1	
		/protein_id="AAB46345.1"	
		/db_xref="GI:1669370"	
	exon	/translation="MFYFPGCGIIFCPQGVVYOIGDVVSVIDEODGKPYAQIRGF IODQEKSAALTWLIPILSPRDPGPASIIIGEDDLPRKMVEYLFVCVAHSEYERK SRSSPPFPVPREPKGYIMTWHPGPTAITTKESANHL"	
		complement(51576 .51756)	
		/gene="WUSC:H.RG083M05.2"	
		/note="Grail prediction, score = 86"	
		/evidence=not_experimental	
	repeat_region	complement(52052 .52329)	
	misc_feature	/rpt_family="J1"	
		55557 .55843	
		/note="match to human EST M79192 (NID:g273505) base 2-289"	
QY	Query Match	100.0%;	Score 201; DB 9; Length 56093;
Db	Best Local Similarity	100.0%;	Pred. No. 1.2e-51;
	Matches 201; Conservative	0; Mismatches	.0; Indels 0; Gaps 0;
QY	1 CAACCTAGTGCAGACATCACCTCCTTAGCCAAATATACACAAGTTCTTAATAACATTACA	60	
Db	35500 CAACTTACTGTGAGACATCACCTCCTTAGCCAAAATATCAACAAGTTCTTAATAACATTACA	35559	
QY	61 AGAACCATATCCCTAGAGAAGGGGAAGAAAGCACTATTCACCCTGTGCATGATATTGT	120	
Db	35560 AGGAACCTATCCCTAGAGAAGGGGAAGAAAGCACTATTCACCCTGTGCATGATATTGT	35619	
QY	121 CAAGTCCCTGCCCTTAATTCGCCATCCCTGATATCATCTCTGGGAAGGACCCCTACCCAGT	180	
Db	35620 CAAGTCCCTGCCCTTAATTCGCCATCCCTGATATCATCTCTGGGAAGGACCCCTACCCAGT	35679	
QY	181 CATTTATATCACCCACACTGC	201	

Db 35680 CATTIATCTACCCCACTGC 35700

```

|||||
RESULT 9
AC007566/c 149194 bp DNA linear PRI 01-MAR-2002
LOCUS AC007566
DEFINITION Homo sapiens BAC clone CTB-10G5 from 7q21-7q22, complete sequence.
ACCESSION AC007566
VERSION AC007566.2 GI:11181861
KEYWORDS HTG.
SOURCE Homo sapiens.
ORGANISM Homo sapiens.
Mammalia; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 149194)
Sulston, J.E. and Waterston, R.
Toward a complete human genome sequence
Genome Res. 8 (11), 1097-1108 (1998)
JOURNAL MEDLINE
PUBMED 99063792
2 (bases 1 to 149194)
Du, Z.
The sequence of Homo sapiens BAC clone CTB-10G5
unpublished (2001)
3 (bases 1 to 149194)
Waterston, R.H.
Direct Submission
Submitted (15-MAY-1999) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
4 (bases 1 to 149194)
Waterston, R.
Direct Submission
Submitted (02-OCT-2000) Department of Genetics, Washington
University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
5 (bases 1 to 149194)
Waterston, R.H.
Direct Submission
Submitted (16-NOV-2000) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
6 (bases 1 to 149194)
Waterston, R.H.
Direct Submission
Submitted (03-JAN-2002) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
7 (bases 1 to 149194)
Waterston, R.H.
Direct Submission
Submitted (06-FEB-2002) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
8 (bases 1 to 149194)
Waterston, R.
Direct Submission
Submitted (01-MAR-2002) Department of Genetics, Washington
University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
On Nov 16, 2000 this sequence version replaced g1:4835815.
-- Genome Center
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: http://genome.wustl.edu/gsc
Contact: saplens@wustl.wustl.edu
----- Summary Statistics
Center Project name: H.RG010G05

```

NOTICE: This sequence may not represent the entire insert of this clone. It may be shorter because we only sequence overlapping clone sections once, or longer because we provide a small overlap between neighboring data submissions.

This sequence was finished as follows unless otherwise noted: all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

MAPPING INFORMATION:

The sequence of this clone was established as part of a mapping and sequencing collaboration between the NIGRI Chromosome 7 Mapping Project (Eric D. Green, Director), John D. McPherson in the Department of Genetics (Washington University), and the Washington University Genome Sequencing Center. For additional information about the map position of this sequence, see <http://www.nhgri.nih.gov/DIR/CTB/CHR7>, send mailto:egreen@nhgri.nih.gov, or see <http://genome.wustl.edu/gsc>

SOURCE INFORMATION:

Clone CTB-10G5 is from the first release of the human BAC library CTB-978SK-B. The library contains cloned DNA from the male fibroblast cell line 978SK. See: Shizuya et al., Proc. Natl. Acad. Sci. USA 89:8794-7 (1992); U-J. Kim et al., Genomics 34:213-8 (1996). This clone is available from Research Genetics, Inc. (<http://www.resgen.com>).

VECTOR: pBelobAC11

Selection: chloramphenicol

NEIGHBORING SEQUENCE INFORMATION:

The clone sequenced to the left is RP5-911H5, 200 base pair overlap. Actual start of this clone is at base position 195 of CTB-10G5; actual end is at base position 150532 of CTB-10G5.

FEATURES
The clone CTB-10G5 contains the entire sequence of CTB-83M5.
Location/Qualifiers

Source

```

1..149194
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="7"
/map="7q21-7q22"
/clone="CTB-10G5"
/clone_1ib="CTB-978SK-B"
1..1634
/rpt_family="L1"
2248..2387
/note="match to EST BG752883 (NID:g14063536)"
2248..2387
/note="match to EST AA149693 (NID:g1720635) zn99406.r1"
2248..2387
/note="match to EST AW579261 (NID:g7254310)"
2248..2387
/note="match to EST BG766882 (NID:g14077535)"
2248..2387
/note="match to EST BI160365 (NID:g14620366)"
2248..2374
/note="match to EST Homo sapiens EST BF758865 (NID:g12106765)"
2248..2287
/note="match to EST AV686676 (NID:g10288539)"
2253..2387
/note="match to EST AV123510 (NID:g10948226)"
2344..2387
/note="match to EST AV686676 (NID:g10288539)"
2696..3066
/rpt_family="MALR"
3108..3392
/rpt_family="Alu"
3540..3628
/note="match to EST AA425526 (NID:g2106267) zw48b03.r1"
3540..3628
/note="similar to Mus musculus EST BB253526 (NID:g8946272)"

```

misc_feature	3542..3632 /note="match to EST AV686676 (NID:g10288539) "
misc_feature	3542..3628 /note="match to EST AA149693 (NID:g1720635) zn99d06.r1"
misc_feature	3542..3628 /note="match to EST AU123510 (NID:g109485226) "
misc_feature	3542..3628 /note="match to EST BG766882 (NID:g14077535) "
misc_feature	3542..3628 /note="match to EST B1160365 (NID:g14620366) "
misc_feature	3542..3609 /note="match to EST AW579261 (NID:g7254310) "
misc_feature	3562..3628 /note="match to EST BE272564 (NID:g9146913) "
misc_feature	3598..3628 /note="match to EST BE299708 (NID:g9183456) "
misc_feature	3714..3785 /note="match to EST BG260659 (NID:g12770475) "
misc_feature	3717..3785 /note="similar to Mus musculus EST BE594936 (NID:g10678674) "
misc_feature	3719..3785 /note="match to EST AA149693 (NID:g1720635) zn99d06.r1"
misc_feature	3719..3785 /note="match to EST AA425526 (NID:g2106267) zw48b03.r1"
misc_feature	3719..3785 /note="match to EST AU123510 (NID:g109485226) "
misc_feature	3719..3785 /note="match to EST BE272564 (NID:g9146913) "
misc_feature	3719..3785 /note="match to EST BE299708 (NID:g9183456) "
misc_feature	3719..3785 /note="match to EST BG766882 (NID:g14077535) "
misc_feature	3719..3785 /note="match to EST B1160365 (NID:g14620366) "
misc_feature	3719..3785 /note="similar to Mus musculus EST BE253526 (NID:g8946272) "
misc_feature	3860..3862 /note="match to EST AV686676 (NID:g10288539) "
repeat_region	4730..4827 /rpt_family="I2" /rpt_family="L1" /rpt_family="LI" /note="match to EST AA425526 (NID:g2106267) zw48b03.r1"
repeat_region	4834..4925 /rpt_family="LI" /note="match to EST AU123510 (NID:g109485226) "
misc_feature	5715..5860 /note="match to EST BE272564 (NID:g9146913) "
misc_feature	5715..5860 /note="match to EST BE299708 (NID:g9183456) "
misc_feature	5715..5860 /note="match to EST BG260659 (NID:g12770475) "
misc_feature	5715..5860 /note="match to EST BG766882 (NID:g14077535) "
misc_feature	5715..5860 /note="match to EST B1160365 (NID:g14620366) "
misc_feature	5715..5860 /note="similar to Mus musculus EST BE594936

QY	121	CAAGTCCTTCCTCCCTAATTCCTCCATTCCTAGATACATCTCGGGAAGACCTACCACGT	180
Db	86384	CAAGTCCTTCCTCCCTAATTCCTCCATTCCTAGATACATCTCGGGAAGACCTACCACGT	86325
QY	181	CATTTATCTACCCCACTGC	201
Db	86324	CATTTATCTACCCCACTGC	86304
RESULT	10		
LOCUS	AA000966	7582 bp	DNA
DEFINITION	AX000966	Sequence 11 from Patent WO9092696.	linear
ACCESSION	AX000966		
VERSION	AX000966.1	GI:7241208	
KEYWORDS			
SOURCE		unidentified.	
ORGANISM		unidentified	
		unclassified.	
REFERENCE		1 (bases 1 to 7582)	
AUTHORS		Beseme, F. and Blond, J.	
TITLE		ENDOGENETIC RETROVIRAL SEQUENCES ASSOCIATED WITH AUTOIMMUNE	
JOURNAL		DISEASES OR WITH PREGNANCY DISORDERS	
		Patent: WO 902696-A 11 21-JAN-1999;	
		BIO MERIEUX (FR); BESEME FREDERIC (FR)	
FEATURES		Location/Qualifiers	
source		1..7582	
		/organism="unidentified"	
		/db_xref="taxon:32644"	
BASE COUNT	2156 a	1876 c	1538 g
ORIGIN		1796 t	216 others

	Query Match	Similarity	100.0%	Score	201	DB	9	Length	149194
	Best	Local	Similarity	100.0%	Pred.	No.	1,3e-51		
	Matches	201	Conservative	0	Mismatches	0	Indels	0	Gaps
QY	1	CAACTTACTTCGACAGCATCACCCTCCTTGAAGCCAAATCAACAAGTCTCTTAAACAATTACA	60						
Db	86504	CAACTTACTTCGACAGCATCACCCTCCTTGAAGCCAAATCAACAAGTCTTAAACAATTACA	86445						
QY	61	AGGAACCTATCCCTGAGAAAGGAGGAAAGAAAGTATTCACCCCTTGGAGATGATATTACT	120						
Db	86444	AGGAACCTATCCCTGAGAAAGGAGGAAAGAAAGTATTCACCCCTTGGAGATGATATTACT	86385						

	Query Match Similarity	99.2%;	Score 199, 4;	DB 6;	Length 7582;	
	Best Local Similarity	98.0%;	Pred. No. 3, 3e-51;			
	Matches 197;	Conservative	4;	Mismatches	0;	Gaps 0;
Oy	1 CAACCTAGTTCGACATCACCCTCCCTTAGCCAAATATACACAAGTTCTTAANAACATTACA	60				
Dd	5201 CAACTAGTAGTGACAGACATCACCTCTGTAGCCAAAATATACACAAGTTCTTAANAACATTACA	5260				
Oy	61 AGGAACCTATCCCTCAGAGAAGGAGAAAAAGACTATTCCACCCTTGTGACATGGTATTAGT	120				
Dd	5261 AGGAACCTATCCCTCAGAGAAGGAGAAAAAGAACTATTCCACCCTTGTGACATGGTATTAGT	5320				
Oy	121 CAAGTCCCTTCCCCTTAATTCGCCCATCCCTAGATACATCTCGAAGAACCCCTACCCAGT	180				
Dd	5321 CAAGTCCCTTCTCTTAATTCGCCCATCCCTAGATACATCTCGAAGAACCCCTACCCAGT	5380				
Oy	181 CATTTATCTACCCCACTGC	201				
	:					
Dd	5381 CATTTATCTACCCCACTGC	5401				
RESULT 11						
LOCUS	AX027480	7582 bp	DNA	linear	PAT 16-SEP-2000	
DEFINITION	Sequence 30 from Patent FR2788784.					
ACCESSION	AX027480					
VERSION	AX027480.1	GI:10188444				
KEYWORDS						
SOURCE	human.					
ORGANISM	Homo sapiens					
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;					
REFERENCE	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo					
AUTHORS	1 (bases 1 to 7582)					
JOURNAL	Mallec,F., Voisset,C. and Paranhos,B.G.					
	Patent: FR 2788784-A 30 28-JUL-2000;					
FEATURES	BIO MERIEUX (FR)					
SOURCE	Location/Qualifiers					
	1..7582					
	/organism="Homo sapiens"					
	/db_xref="taxon:9606"					
BASE COUNT	2156 a 1876 c 1538 g 1796 t	216 others				

ORIGIN

Query Match 99.2%; Score 199.4; DB 6; Length 7582;
Best Local Similarity 98.0%; Pred. No. 3.3e-51;
Matches 197; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 1 CAACTTAGTGGAGACATCACCCTCTTACGCAAAATATCAACAAGTTCTTAAACATTACA 60
DB 5201 CAACTTAGTGGAGACATCACCCTCTTACGCAAAATATCAACAAGTTCTTAAACATTACA 5260
QY 61 AGGAACCTATCCCTGAGAGAGAGGAAAGAAAGAACTATTCACACCTGTGACATGGTATTAGT 120
DB 5261 AGGAACCTATCCCTGAGAGAGAGGAAAGAAAGAACTATTCACACCTGTGACATGGTATTAGT 5320
QY 121 CAAGTCCCTCTCTCTAATATCCCATCCCTAGATCATCCTGGAGAGACCCCTACCCAGT 180
DB 5321 CAAGTCCCTCTCTCTAATATCCCATCCCTAGATCATCCTGGAGAGACCCCTACCCAGT 5380
QY 181 CATTATATCTACCCCAACTGC 201
DB 5381 CATTATATCTACCCCAACTGC 5401

RESULT 12

AX000959 2006 bp DNA linear PAT 10-MAR-2000
LOCUS AX000959 Sequence 4 from Patent WO9902696.
DEFINITION AX000959
ACCESSION AX000959.1 GI:7241201
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES
source

1 (bases 1 to 2006)
Beseme, F. and Blond, J.
ENDGENETIC RETROVIRAL SEQUENCES, ASSOCIATED WITH AUTOIMMUNE
DISEASES OR WITH PREGNANCY DISORDERS
Patent: WO 9902696-A 4 21-JAN-1999;
BIO MERIEUX (FR); BESEME FREDERIC (FR)
Location/Qualifiers
1. 2006
/organism="unidentified"
/db_xref="taxon:32644"

BASE COUNT 625 a 532 c 373 g 475 t 1 others
ORIGIN

Query Match 98.4%; Score 197.8; DB 6; Length 2006;
Best Local Similarity 99.0%; Pred. No. 9.2e-51;
Matches 199; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CAACTTAGTGGAGACATCACCCTCTTACGCAAAATATCAACAAGTTCTTAAACATTACA 60
DB 1042 CAACTTAGTGGAGACATCACCCTCTTACGCAAAATATCAACAAGTTCTTAAACATTACA 1101
QY 61 AGGAACCTATCCCTGAGAGAGAGGAAAGAAAGAACTATTCACACCTGTGACATGGTATTAGT 120
DB 1102 AGGAACCTATCCCTGAGAGAGAGGAAAGAAAGAACTATTCACACCTGTGACATGGTATTAGT 1161
QY 121 CAAGTCCCTCTCTCTAATATCCCATCCCTAGATCATCCTGGAGAGACCCCTACCCAGT 180
DB 1162 CAAGTCCCTCTCTCTAATATCCCATCCCTAGATCATCCTGGAGAGACCCCTACCCAGT 1221
QY 181 CATTATATCTACCCCAACTGC 201
DB 1222 CATTATATCTACCCCAACTGC 1242

RESULT 13
AX027473 2006 bp DNA linear PAT 16-SEP-2000
LOCUS AX027473
DEFINITION Sequence 23 from Patent FR2788784.
ACCESSION AX027473
VERSION AX027473.1 GI:10188437

KEYWORDS

SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

REFERENCE
1 (bases 1 to 2006)
Mallet, F., Voisset, C. and Paranhos, B. G.
Patent: FR 2788784-A 23 28-JUN-2000;
BIO MERIEUX (FR)
Location/Qualifiers
1. 2006
/organism="Homo sapiens"
/db_xref="taxon:9606"

FEATURES
source

BASE COUNT 625 a 532 c 373 g 475 t 1 others
ORIGIN

Query Match 98.4%; Score 197.8; DB 6; Length 2006;
Best Local Similarity 99.0%; Pred. No. 9.2e-51;
Matches 199; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CAACTTAGTGGAGACATCACCCTCTTACGCAAAATATCAACAAGTTCTTAAACATTACA 60
DB 1042 CAACTTAGTGGAGACATCACCCTCTTACGCAAAATATCAACAAGTTCTTAAACATTACA 1101
QY 61 AGGAACCTATCCCTGAGAGAGAGGAAAGAAAGAACTATTCACACCTGTGACATGGTATTAGT 120
DB 1102 AGGAACCTATCCCTGAGAGAGAGGAAAGAAAGAACTATTCACACCTGTGACATGGTATTAGT 1161
QY 121 CAAGTCCCTCTCTCTAATATCCCATCCCTAGATCATCCTGGAGAGACCCCTACCCAGT 180
DB 1162 CAAGTCCCTCTCTCTAATATCCCATCCCTAGATCATCCTGGAGAGACCCCTACCCAGT 1221
QY 181 CATTATATCTACCCCAACTGC 201
DB 1222 CATTATATCTACCCCAACTGC 1242

RESULT 14

AF072503 2006 bp mRNA linear PRI 10-FEB-1999
LOCUS AF072503
DEFINITION Homo sapiens endogenous retrovirus W sequence.
ACCESSION AF072503
VERSION AF072503.1 GI:4262287
KEYWORDS
SOURCE
ORGANISM
Homo sapiens

REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES
source

1 (bases 1 to 2006)
Blond, J. L., Beseme, F., Duret, L., Boulon, O., Bedin, F., Perron, H.,
Mandrand, B. and Mallet, F.
Molecular characterization and placental expression of HERV-W, a
new human endogenous retrovirus family
J. Virol. 73 (2), 1175-1185 (1999)
2 (bases 1 to 2006)
Blond, J. L., Beseme, F. and Mallet, F.
Direct Submission
Submitted (19-JUN-1998) UN103 CNRS-biomerieux, ENS Lyon, 46 allée
d'Italie, Lyon, Cedex 07 69364, France
Location/Qualifiers
1. 2006
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="cl.p122"
/tissue_type="placenta"
/clone_lib="Clontech 5"-stretch plus library, Cat number
HU5014a"
1. 2006
/note="corresponds to pol and env"
1. 2006
/rpt_family="HERV-W"

misc_feature
repeat_region

Search completed: May 2, 2003, 15:26:45
Job time : 759 secs

XX Claim 3; Page 142-145; 225pp; French.
PS
XX
XX The present invention relates to new nucleic acid sequences of human
CC endogenous retrovirus, HERV-7g, which is located on chromosome 7g.
CC Regulatory elements associated with HERV-7g may alter expression of other
CC genes (even remote genes) on the same chromosome, inducing immunological
CC and/or neurological changes (which may be pathological or protective/
CC curative). HERV-7g peptides can be used to improve efficiency of the
CC immune response, e.g. in immunotherapy. HERV-7g peptides and their coding
CC sequences can be used in immunogenic or vaccinating compositions, for
CC protection against autoimmune diseases, particularly multiple sclerosis.
CC The peptides may also be used (by sequence comparison) to detect/identify
CC endogenous retroviruses that are abnormally expressed in cancer,
CC neuropathologies or other autoimmune diseases. The present sequence was
CC used to illustrate the invention.
XX
SQ Sequence 2055 BP; 576 A; 574 C; 376 G; 529 T; 0 other;
Query Match 100.0%; Score 201; DB 21; Length 2055;
Best Local Similarity 100.0%; Pred. No. 1.3e-54;
Matches 201; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CAACCTAGTTGAGACATCACCTCCTTACGCCAAATATCAAGTCTTAAACATTACA 60
Db 12 CAACCTAGTTGAGACATCACCTCCTTACGCCAAATATCAAGTCTTAAACATTACA 71
QY 61 AGGAACCTATCCCTGAGAGAGGAAAGAACTATTCACCTCTGTGACATGTAATTAGT 120
Db 72 AGGAACCTATCCCTGAGAGAGGAAAGAACTATTCACCTCTGTGACATGTAATTAGT 131
QY 121 CAAGTCCCTCCCTCTAATTTCCCATCCTAGATACATCCTGTGGAGGAGCCTTACCAGT 180
Db 132 CAAGTCCCTCCCTCTAATTTCCCATCCTAGATACATCCTGTGGAGGAGCCTTACCAGT 191
QY 181 CATTTATCTACCCCACTGC 201
Db 192 CATTTATCTACCCCACTGC 212
RESULT 2
ABN97927
ID ABN97927 standard; DNA; 2599 BP.
XX
XX ABN97927;
AC
XX
XX 01-AUG-2002 (first entry)
DT
XX
XX Human retroviral sequence HERV 7 env.
DE
XX
XX Autoimmune disease; HERV-7g; chromosome 7g; immunotherapy;
KW multiple sclerosis; ds.
XX
XX Human retrovirus.
OS
XX
XX WO9967395-A1.
FN
XX
XX 29-DEC-1999.
PD
XX
XX 23-JUN-1999; 99WO-FR01513.
PF
XX
XX 23-JUN-1998; 98FR-0007920.
PR
XX
XX (INRM) INSERM INST NAT SANTE & RECH MEDICALE.
PA
XX
XX Alliel PM, Berlin J, Rieger F;
PI
XX
XX WPI; 2000-160587/14.
DR
XX
XX New nucleic acid sequences of human endogenous retrovirus, HERV-7g,
PT used for diagnosis, treatment and prevention of autoimmune and
PT neurological diseases
XX

PS Claim 1; Page 128-129; 225pp; French.
XX
XX
XX The present invention relates to new nucleic acid sequences of human
CC endogenous retrovirus, HERV-7g, which is located on chromosome 7g.
CC Regulatory elements associated with HERV-7g may alter expression of other
CC genes (even remote genes) on the same chromosome, inducing immunological
CC and/or neurological changes (which may be pathological or protective/
CC curative). HERV-7g peptides can be used to improve efficiency of the
CC immune response, e.g. in immunotherapy. HERV-7g peptides and their coding
CC sequences can be used in immunogenic or vaccinating compositions, for
CC protection against autoimmune diseases, particularly multiple sclerosis.
CC The peptides may also be used (by sequence comparison) to detect/identify
CC endogenous retroviruses that are abnormally expressed in cancer,
CC neuropathologies or other autoimmune diseases. The present sequence was
CC used to illustrate the invention.
XX
SQ Sequence 2599 BP; 744 A; 718 C; 495 G; 642 T; 0 other;
Query Match 100.0%; Score 201; DB 21; Length 2599;
Best Local Similarity 100.0%; Pred. No. 1.4e-54;
Matches 201; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CAACCTAGTTGAGACATCACCTCCTTACGCCAAATATCAAGTCTTAAACATTACA 60
Db 550 CAACCTAGTTGAGACATCACCTCCTTACGCCAAATATCAAGTCTTAAACATTACA 609
QY 61 AGGAACCTATCCCTGAGAGAGGAAAGAACTATTCACCTCTGTGACATGTAATTAGT 120
Db 610 AGGAACCTATCCCTGAGAGAGGAAAGAACTATTCACCTCTGTGACATGTAATTAGT 669
QY 121 CAAGTCCCTCCCTCTAATTTCCCATCCTAGATACATCCTGTGGAGGAGCCTTACCAGT 180
Db 670 CAAGTCCCTCCCTCTAATTTCCCATCCTAGATACATCCTGTGGAGGAGCCTTACCAGT 729
QY 181 CATTTATCTACCCCACTGC 201
Db 730 CATTTATCTACCCCACTGC 750
RESULT 3
AAD24195
ID AAD24195 standard; CDNA; 2930 BP.
XX
XX AAD24195;
AC
XX
XX 07-MAY-2002 (first entry)
DT
XX
XX Human syncytin cDNA.
DE
XX
XX Human; syncytin; pre-eclampsia; gestational trophoblast disorder;
KW choriocarcinoma; hydatiform mole; placental site tumour; abortion;
KW envelope gene; human endogenous defective retrovirus; HERV-W, ss.
XX
XX Homo sapiens.
OS
XX
XX Key Location/Qualifiers
FH 930..2546
FT /*tag= a
FT /product= "Syncytin"
FN
XX
XX WO200204678-A2.
FN
XX
XX 17-JAN-2002.
PD
XX
XX 09-JUL-2001; 2001WO-US21719.
PE
XX
XX 07-JUL-2000; 2000US-216657P.
PR
XX
XX (GENY) GENETICS INST INC.
PA
XX
XX Keith JC, McCoy JM, M1 S;
PI
XX
XX WPI; 2002-171727/22.
DR

DR P-PSDB: AAE14540.

XX Identifying a compound for treating a subject with or at risk of

PT developing preeclampsia, comprises determining whether the expression

PT or activity of syncytin in the cell is modulated in the presence of a

PT test compound

XX

PS Disclosure; Page 39-42; 43pp; English.

XX

CC The invention relates to identifying compounds which are modulators

CC of syncytin expression. The syncytin modulators are useful in diagnosis

CC and treatment of preeclampsia and gestational trophoblast disorders (e.g.

CC choriocarcinoma, hydatiform mole, placental site tumour and missed/

CC incomplete abortion). Syncytin is a human gene derived from the

CC envelope gene of human endogenous defective retrovirus, HERV-W. The

CC present invention is based partly on the discovery that syncytin

CC expression is dramatically reduced in preeclampsia, and is also

CC mis-localised to the apical syncytiotrophoblast membrane. The present

CC sequence is human syncytin cDNA.

XX

SO Sequence 2930 BP; 842 A; 800 C; 571 G; 717 T; 0 other;

Query Match 100.0%; Score 201; DB 24; Length 2930;

Best Local Similarity 100.0%; Pred. No. 1.5e-54;

Matches 201; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 CAACTAGTTCGACATCACCCTCTTACCAATATCAACAAGTCTTAAACATTACA 60

DB 551 CAACCTAGTTCGACATCACCCTCTTACCAATATCAACAAGTCTTAAACATTACA 610

OY 61 AGAACCCTATCCCTGAGAGAGGAAAGAACTATTCACCTTGTGACATGTTTGT 120

DB 611 AGAACCCTATCCCTGAGAGAGGAAAGAACTATTCACCTTGTGACATGTTTGT 670

OY 121 CAAGTCCCTTCCTTAATTCCTCCATCCCTAGATATCTGGAAGACCTACCAGT 180

DB 671 CAAGTCCCTTCCTTAATTCCTCCATCCCTAGATATCTGGAAGACCTACCAGT 730

OY 181 CATTTATCTACCCCACTGC 201

DB 731 CATTTATCTACCCCACTGC 751

RESULT 4

AAZ7526

ID AAZ7526 standard; cDNA; 2946 BP.

XX

AC AAZ7526;

XX

DT 10-AUG-1999 (first entry)

XX

DE Human secreted protein AJ172_2 cDNA.

XX

KW Secreted protein; testes; brain; blood; placenta; human; murine; thymus;

KW bone marrow; treatment; prevention; nutrition; cytokine; immune; vaccine;

KW cell proliferation; cell differentiation; suppressor; tumour inhibition;

KW haematopoiesis regulator; activin; inhibin; chemotactic; chemokineic;

KW haemostatic; thrombolytic; receptor; ligand; anti-inflammatory; tumour;

KW cadherin; tumour invasion suppressor; gene therapy; tissue growth; ss.

XX

OS Homo sapiens.

XX

PN WO9926972-A1.

XX

PD 03-JUN-1999.

XX

XX 17-NOV-1998; 98WO-US24614.

XX

PR 20-OCT-1998; 98US-0175928.

XX

PR 21-NOV-1997; 97US-0976110.

XX

PR 18-MAY-1998; 98US-0080478.

XX

PA (GENY) GENETICS INST INC.

XX

PI Collins-Racie LA, Evans C, Jacobs K, Lavallie ER;

PI McCoy JM, Meberg D, Treacy M;

XX

DR WPI: 1999-357813/30.

DR P-PSDB: AAY08622.

XX

PT New polynucleotides encoding secreted proteins

XX

PS Claim 13a; Page 100-101; 142pp; English.

XX

CC This invention describes novel human secreted proteins encoded by

CC polynucleotides isolated from human adult testes, adult brain, adult

CC blood or adult placenta, or murine adult bone marrow or thymus cDNA

CC libraries. The products of the invention are predicted to have biological

CC activities which would make them suitable for treating, preventing or

CC ameliorating medical conditions in humans and animals, although no

CC supporting data is given. Suggested activities include nutritional

CC activity, cytokine and cell proliferation/differentiation activity,

CC immune stimulating (e.g. as vaccines) or suppressing activity,

CC haematopoiesis regulating activity, tissue growth activity,

CC activin/inhibin activity, chemotactic/chemokineic activity, haemostatic

CC and thrombolytic activity, receptor/ligand activity, anti-inflammatory

CC activity, cadherin/tumour invasion suppressor activity, and tumour

CC inhibition activity. The polynucleotides are also stated to be useful

CC for gene therapy.

XX

SO Sequence 2946 BP; 858 A; 801 C; 570 G; 717 T; 0 other;

Query Match 100.0%; Score 201; DB 20; Length 2946;

Best Local Similarity 100.0%; Pred. No. 1.5e-54;

Matches 201; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 CAACTAGTTCGACATCACCCTCTTACCAATATCAACAAGTCTTAAACATTACA 60

DB 549 CAACCTAGTTCGACATCACCCTCTTACCAATATCAACAAGTCTTAAACATTACA 608

OY 61 AGAACCCTATCCCTGAGAGAGGAAAGAACTATTCACCTTGTGACATGTTTGT 120

DB 609 AGAACCCTATCCCTGAGAGAGGAAAGAACTATTCACCTTGTGACATGTTTGT 668

OY 121 CAAGTCCCTTCCTTAATTCCTCCATCCCTAGATATCTGGAAGACCTACCAGT 180

DB 669 CAAGTCCCTTCCTTAATTCCTCCATCCCTAGATATCTGGAAGACCTACCAGT 728

OY 181 CATTTATCTACCCCACTGC 201

DB 729 CATTTATCTACCCCACTGC 749

RESULT 5

AAZ59468

ID AAZ59468 standard; cDNA; 2946 BP.

XX

AC AAZ59468;

XX

DT 11-APR-2000 (first entry)

XX

DE Human secreted protein AJ172_2 polynucleotide sequence.

XX

KW Human; secreted protein; disease diagnosis; pre-eclampsia; cancer;

KW placental pathology; metastasis inhibition; nutritional activity;

KW immune stimulator; haematopoiesis regulator; tissue growth;

KW tumour inhibitor; anti-inflammatory; clone AJ172_2; ATCC_98115;

KW gene therapy; ss.

XX

OS Homo sapiens.

XX

PN WO9960020-A1.

XX

PD 25-NOV-1999.

XX

PF 17-MAY-1999; 99WO-US10915.

XX 18-MAY-1998; 98US-0080478.
PR 20-OCT-1998; 98US-0175928.
XX
XX (GEMV) GENETICS INST INC.
PI Jacobs K, McCoy JM, LaVallie ER, Collins-Racie LA, Evans C;
PI Werberg D, Ml S, Treacy M;
DR WPI: 2000-116311/10.
DR P-PSDB: AAY67313.
XX
PT New polynucleotides encoding secreted cDNA libraries, used to develop
PT products for the diagnosis and treatment of neoplastic disease
XX
PS Claim 14; Page 107-108; 149pp; English.
XX
CC This is the human secreted protein AJ172_2 nucleotide sequence, obtained
CC from a human adult testes cDNA library. The invention relates to secreted
CC human and murine proteins. The polynucleotides and proteins are predicted
CC to have biological activities which would make them suitable for
CC treating, preventing or ameliorating medical conditions in humans and
CC animals. Detection of the levels of the proteins can be used for the
CC diagnosis of e.g. pre-eclampsia, placental pathology or cancer. Agents
CC which modulate the expression or function of the proteins may be used for
CC treating a neoplastic disease and inhibiting metastasis. Other suggested
CC activities include nutritional activity (e.g. in feeds), cytokine and
CC cell proliferation/differentiation activity, immune stimulating
CC (e.g. as vaccines) or suppressing activity, haematopoiesis regulating
CC activity, tissue growth activity, activin/inhibin activity,
CC chemotactic/chemokinetic activity, haemostatic and thrombolytic activity,
CC receptor/ligand activity, anti-inflammatory activity, cadherin/tumour
CC invasion suppressor activity, and tumour inhibition activity. The
CC polynucleotide sequences are also stated to be useful for gene therapy.
XX
SQ Sequence 2946 BP; 858 A; 801 C; 570 G; 717 T; 0 other:
Query Match 100.0%; Score 201; DB 21; Length 2946;
Best Local Similarity 100.0%; Pred. No. 1.5e-54;
Matches 201; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CAACCTAGTGGACAGATCCCTGCTAGCAATATCAACAGTCTTAAACATTACA 60
DB 549 CAACCTAGTGGACAGATCCCTGCTAGCAATATCAACAGTCTTAAACATTACA 608
QY 61 AGAACCTATCCCTGAGAAGAGGAAAGAACTATTCCACCCTTGACATGCTATTAGT 120
DB 609 AGAACCTATCCCTGAGAAGAGGAAAGAACTATTCCACCCTTGACATGCTATTAGT 668
QY 121 CAAGTCCTTCCCTCTTAATTCCTCATTCCTAGATACATCTGGGAAGAGACCTACCCAGT 180
DB 669 CAAGTCCTTCCCTCTTAATTCCTCATTCCTAGATACATCTGGGAAGAGACCTACCCAGT 728
QY 181 CATTTATCTACCCCAACTGC 201
DB 729 CATTTATCTACCCCAACTGC 749
RESULT 6
ABN97929
ID ABN97929 standard; DNA; 10499 BP.
XX
XX ABN97929;
AC
XX
XX 01-AUG-2002 (first entry)
DT
XX
XX Human retroviral sequence HERV-7q.
DE
XX
XX Autimmune disease; HERV-7q; chromosome 7q; immunotherapy;
KW multiple sclerosis; ds.
XX
XX Human retrovirus.
OS

PN WO967395-A1.
XX
PD 29-DEC-1999.
XX
XX 23-JUN-1999; 99WO-FR01513.
PF
XX 23-JUN-1998; 98FR-0007920.
PR
XX (INRM) INSERM INST NAT SANTE & RECH MEDICALE.
PA
XX Allel PM, Perlin J, Rieger F;
PI
XX WPI: 2000-160587/14.
DR
XX
PT New nucleic acid sequences of human endogenous retrovirus, HERV-7q,
PT used for diagnosis, treatment and prevention of autoimmune and
PT neurological diseases
XX
PS Claim 3; Fig 1; 225pp; French.
XX
XX The present invention relates to new nucleic acid sequences of human
XX endogenous retrovirus, HERV-7q, which is located on chromosome 7q.
XX Regulatory elements associated with HERV-7q may alter expression of other
XX genes (even remote genes) on the same chromosome, inducing immunological
XX and/or neurological changes (which may be pathological or protective/
XX curative). HERV-7q peptides can be used to improve efficiency of the
XX immune response, e.g. in immunotherapy. HERV-7q peptides and their coding
XX sequences can be used in immunogenic or vaccinating compositions, for
XX protection against autoimmune diseases, particularly multiple sclerosis.
XX The peptides may also be used (by sequence comparison) to detect/identify
XX endogenous retroviruses that are abnormally expressed in cancer,
XX neuropathologies or other autoimmune diseases. The present sequence was
XX used to illustrate the invention.
SQ Sequence 10499 BP; 3048 A; 2676 C; 2280 G; 2495 T; 0 other:
Query Match 100.0%; Score 201; DB 21; Length 10499;
Best Local Similarity 100.0%; Pred. No. 2.3e-54;
Matches 201; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CAACCTAGTGGACAGATCCCTGCTAGCAATATCAACAGTCTTAAACATTACA 60
DB 7500 CAACCTAGTGGACAGATCCCTGCTAGCAATATCAACAGTCTTAAACATTACA 7559
QY 61 AGAACCTATCCCTGAGAAGAGGAAAGAACTATTCCACCCTTGACATGCTATTAGT 120
DB 7560 AGAACCTATCCCTGAGAAGAGGAAAGAACTATTCCACCCTTGACATGCTATTAGT 7619
QY 121 CAAGTCCTTCCCTCTTAATTCCTCATTCCTAGATACATCTGGGAAGAGACCTACCCAGT 180
DB 7620 CAAGTCCTTCCCTCTTAATTCCTCATTCCTAGATACATCTGGGAAGAGACCTACCCAGT 7679
QY 181 CATTTATCTACCCCAACTGC 201
DB 7680 CATTTATCTACCCCAACTGC 7700
RESULT 7
ABL61744
ID ABL61744 standard; DNA; 56093 BP.
XX
XX ABL61744;
AC
XX
XX 15-MAY-2002 (first entry)
DT
XX
XX Colon adenocarcinoma related gene sequence SEQ ID NO:81.
DE
XX
XX Human; cancer; colon; breast; ovary; oesophagus; kidney; thyroid;
KW stomach; lung; prostate; pancreas; carcinoma; antitumour; cancerous;
KW cytostatic; gene therapy; antineoplastic; Wilm's tumour; adenocarcinoma;
KW gene; ds.
XX
XX Homo sapiens.
OS

XX WO200194629-A2.
PN
XX
PD 13-DEC-2001.
XX
PF 30-MAY-2001; 2001WO-US10838.
XX
PR 05-JUN-2000; 2000US-209473P.
PR 05-JUN-2000; 2000US-209531P.
PR 18-SEP-2000; 2000US-233133P.
PR 18-SEP-2000; 2000US-233617P.
PR 20-SEP-2000; 2000US-234009P.
PR 20-SEP-2000; 2000US-234034P.
PR 20-SEP-2000; 2000US-234052P.
PR 22-SEP-2000; 2000US-234509P.
PR 22-SEP-2000; 2000US-234567P.
PR 25-SEP-2000; 2000US-234923P.
PR 25-SEP-2000; 2000US-234924P.
PR 25-SEP-2000; 2000US-235077P.
PR 25-SEP-2000; 2000US-235082P.
PR 25-SEP-2000; 2000US-235134P.
PR 25-SEP-2000; 2000US-235280P.
PR 26-SEP-2000; 2000US-235637P.
PR 26-SEP-2000; 2000US-235638P.
PR 27-SEP-2000; 2000US-235711P.
PR 27-SEP-2000; 2000US-235720P.
PR 27-SEP-2000; 2000US-235840P.
PR 27-SEP-2000; 2000US-235863P.
PR 28-SEP-2000; 2000US-236028P.
PR 28-SEP-2000; 2000US-236033P.
PR 28-SEP-2000; 2000US-236033P.
PR 28-SEP-2000; 2000US-236034P.
PR 28-SEP-2000; 2000US-236109P.
PR 28-SEP-2000; 2000US-236111P.
PR 29-SEP-2000; 2000US-236894P.
PR 29-SEP-2000; 2000US-236894P.
PR 02-OCT-2000; 2000US-237172P.
PR 02-OCT-2000; 2000US-237173P.
PR 02-OCT-2000; 2000US-237278P.
PR 02-OCT-2000; 2000US-237294P.
PR 02-OCT-2000; 2000US-237295P.
PR 02-OCT-2000; 2000US-237315P.
PR 03-OCT-2000; 2000US-237425P.
PR 03-OCT-2000; 2000US-237596P.
PR 03-OCT-2000; 2000US-237604P.
PR 03-OCT-2000; 2000US-237606P.
PR 03-OCT-2000; 2000US-237608P.
PR 01-NOV-2000; 2000US-244867P.
PR 01-NOV-2000; 2000US-245084P.
XX
PA (AVAL-) AVALON PHARM.
XX
PI Young PE, Augustus M, Carter KC, Ebner R, Endress G, Horrigan S;
PI Soppet DR, Weaver Z;
XX
DR WPI; 2002-186264/24.
XX
XX
PT Screening for anti-neoplastic agent involves exposing cells to a
PT chemical agent to be tested for anti-neoplastic activity, and
PT determining a change in expression of a gene of a signature gene set
XX
XX Claim 1; SEQ ID 81; 44pp; English.
XX
XX The present invention describes a method (M1) for screening for an
XX anti-neoplastic agent. The method involves exposing cells to a chemical
XX agent to be tested for anti-neoplastic activity, determining a change in
XX expression of at least one gene (I) of a signature gene set, where (I)
XX comprises a sequence (S) selected from 8447 sequences (given in ABL61664
XX to ABL70110), or is at least 95% identical to (S), where a change in
XX expression is indicative of anti-neoplastic activity. (I) has cytosol
XX activity and can be used in gene therapy. M1 can be used for screening
XX an anti-neoplastic agent, and can be used for producing a product which
XX is the data collected with respect to the anti-neoplastic agent as a

CC	result of M1, and the data is sufficient to convey the chemical
CC	structure and/or properties of the agent. M1 can be used in the
CC	treatment of cancer such as colon, breast, stomach, lung, thyroid,
CC	oesophageal, ovarian, kidney, prostate or pancreatic cancer,
CC	adenocarcinoma, carcinoma, clear cell cancer, infiltrating ductal cancer,
CC	infiltrating lobular cancer, squamous cell carcinoma, neuroendocrine
CC	carcinoma, papillary carcinoma and Wilms' tumour.
XX	
SQ	Sequence 56093 BP: 16164 A; 12346 C; 10702 G; 16881 T; 0 other;
	Query Match 100.0%; Score 201; DB 24; Length 56093;
	Best Local Similarity 100.0%; Pred No. 4e-54;
	Matches 201; Conservative 0; Mismatches 0; Indels 0; Gaps 0
QY	1 CAACTAGTGTGAGACATCATCCTCCTTAGCCAAATATCAACAAGTTCTTAAACATTACA 60
Db	35500 CAACCTAGTGTGAGACATCATCCTCCTTAGCCAAATATCAACAAGTTCTTAAACATTACA 35559
QY	61 AGGAACCTATCCTCTGAGAGAGGAGGAAAAGACATATTCACCCCTTGTGACATGGTATTAGT 120
Db	35560 AGGAACCTATCCTCTGAGAGAGGAGGAAAAGACATATTCACCCCTTGTGACATGGTATTAGT 35619
QY	121 CAAGTCCTCTCCTCCTTAATTCCTCCATCCTAGATATACATCCTGGGAAGAGACCTACCCAGT 180
Db	35620 CAAGTCCTCTCCTCCTTAATTCCTCCATCCTAGATATACATCCTGGGAAGAGACCTACCCAGT 35679
QY	181 CATTTATCTACCCCAACTGC 201
Db	35680 CATTTATCTACCCCAACTGC 35700
XX	
XX	RESULT 8
AXX	AXX25665
XX	ID AAX25665 standard; cDNA to mRNA; 7582 BP.
AC	AAX25665;
XX	
DT	21-MAY-1999 (first entry)
XX	
DE	Complete human endogenous retrovirus W genome.
XX	
KW	Clone; human endogenous retrovirus; genome; autoimmune disease;
KW	multiple sclerosis; Rheumatoid polyarthritis; insulin-dependent diabetes;
KW	disseminated lupus erythematosus; pregnancy; chromosomal marker; ss.
XX	
OS	Human endogenous retrovirus.
XX	
PN	W09902696-A1.
XX	
PD	21-JAN-1999.
XX	
PF	06-JUL-1998; 98W0-FR01442.
XX	
PR	07-JUL-1997; 97FR-0008815.
XX	
PA	(INMR) BIO MERIEUX.
XX	
PI	Beseme F, Blond JL, Bouton O, Mallet F, Mandrand B;
XX	
DR	WPL; 1999-120897/10.
XX	
PT	New nucleic acid sequences from human endogenous retrovirus-W -
PT	expressed exclusively in placenta and useful in diagnosis and
PT	therapy of autoimmune disease, and abnormal or failed pregnancy
XX	
PS	Claim 1; Page 71-74; 106pp; French.
XX	
CC	This sequence represents the complete sequence of the human endogenous
CC	retrovirus (HERV) W genome. The nucleic acids, their fragments or
CC	peptides encoded by them are markers of autoimmune disease (e.g. multiple
CC	sclerosis, rheumatoid polyarthritis, disseminated lupus erythematosus,
CC	insulin-dependent diabetes and related pathologies) and of abnormal or
CC	unsuccessful pregnancy and can be used as chromosomal markers for

CC susceptibility to these conditions, or proximity markers of genes
CC associated with this susceptibility.

XX Sequence 7582 BP; 2156 A; 1877 C; 1537 G; 1796 T; 2 U; 214 other;

Query Match 99.28; Score 199.4; DB 20; Length 7582;
Best Local Similarity 98.08; Pred. No. 6.8e-54;
Matches 197; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 1 CAACCTTAGTTGACAGACATCACCTCTTAGCTAAATATCAACAAGTCTTAAACATTACA 60
DB 5201 CAACCTTAGTTGACAGACATCACCTCTTAGCTAAATATCAACAAGTCTTAAACATTACA 5260
QY 61 AGGAACCTATCCCTGAGAGAGGAGGAGAAAGAAATATTCACCCCTTGACATGTATTAGT 120
DB 5261 AGGAACCTATCCCTGAGAGAGGAGGAGGAGAAAGAAATATTCACCCCTTGACATGTATTAGT 5320
QY 121 CAAGTCCCTTCCTCTAATTCCTCCATCCCTAATACATCCTGGAGAGACCTACCCAGT 180
DB 5321 CAAGTCCCTTCCTCTAATTCCTCCATCCCTAATACATCCTGGAGAGACCTACCCAGT 5380
QY 181 CATTATATACCCCAACTGC 201
DB 5381 CATTATATACCCCAACTGC 5401

RESULT 9

AAAS9215
ID AAAS9215 standard; DNA: 7582 BP.

XX AC AAAS9215;

XX DT 07-NOV-2000 (first entry)

XX DE Human endogenous retrovirus W (HERV-W) sequence.

XX KW Autoimmune disease; retrovirus; human endogenous retrovirus W; HERV-W;
XX gag gene; pregnancy; multiple sclerosis; T cell proliferation; ss.

XX OS Human endogenous retrovirus.

XX Key Location/Qualifiers

FT LTR 1..120

FT LTR /tag- a

FT LTR /note- "R of 5' LTR"

FT LTR 121..575

FT LTR /tag- b

FT LTR /note- "U5 of 5' LTR"

FT primer_bind 579..596

FT CDS /tag- c

FT CDS 5581..7194

FT CDS /tag- d

FT CDS /note- "ORF1 env538"

FT CDS 7039..7194

FT CDS /tag- e

FT CDS /note- "ORF2 52 AA"

FT CDS 7112..7255

FT CDS /tag- f

FT CDS /note- "ORF3 48 AA"

FT misc_feature 7244..7254

FT LTR /tag- g

FT LTR /note- "polypurine tract."

FT LTR 7256..7582

FT LTR /tag- h

FT LTR /note- "U3-R of 3' LTR"

FT LTR 7563..7569

FT LTR /tag- i

XX polyA_signal

XX WO200043521-A2.

XX 27-JUL-2000.

XX PD 21-JAN-2000; 2000WO-FR00144.

XX PF

XX 21-JAN-1999; 99FR-0000888.

XX (INMR) BIO MERIEUX.

XX PA Paranhos-Baccala G, Mallet F, Voisset C;

XX WPI; 2000-499229/44.

XX PT New nucleic acid from human endogenous retrovirus, useful e.g. for

XX diagnosis of autoimmune disease and complications of pregnancy,

XX PT contains at least part of the gag gene

XX PS disclosure; Page 49-52; 53pp; French.

XX CC The present sequence represents an endogenous retrovirus, which is

XX associated with an autoimmune disease, and is integrated into the human

XX genome. The retrovirus is human endogenous retrovirus W (HERV-W). The

XX HERV-W retrovirus is associated with autoimmune disease, failure of

XX pregnancy or disorders of pregnancy. HERV-W nucleic acid fragments, or

XX proteins derived from it, are useful for diagnosis of autoimmune

XX disease (specifically multiple sclerosis) and for monitoring pregnancy.

XX CC The nucleic acid fragments may also be used for in situ labelling of

XX isolated chromosomes while the transcription product can be used to

XX study or monitor T cell proliferation in vitro.

XX SQ Sequence 7582 BP; 2156 A; 1876 C; 1538 G; 1796 T; 216 other;

Query Match 99.28; Score 199.4; DB 21; Length 7582;

Best Local Similarity 98.08; Pred. No. 6.8e-54;
Matches 197; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 1 CAACCTTAGTTGACAGACATCACCTCTTAGCTAAATATCAACAAGTCTTAAACATTACA 60
DB 5201 CAACCTTAGTTGACAGACATCACCTCTTAGCTAAATATCAACAAGTCTTAAACATTACA 5260
QY 61 AGGAACCTATCCCTGAGAGAGGAGGAGGAGAAAGAACTATTCACCCCTTGACATGTATTAGT 120
DB 5261 AGGAACCTATCCCTGAGAGAGGAGGAGGAGGAGGAGAAAGAACTATTCACCCCTTGACATGTATTAGT 5320
QY 121 CAAGTCCCTTCCTCTAATTCCTCCATCCCTAATACATCCTGGAGAGACCTACCCAGT 180
DB 5321 CAAGTCCCTTCCTCTAATTCCTCCATCCCTAATACATCCTGGAGAGACCTACCCAGT 5380
QY 181 CATTATATACCCCAACTGC 201
DB 5381 CATTATATACCCCAACTGC 5401

RESULT 10

AAK25658
ID AAK25658 standard; cDNA to mRNA: 2006 BP.

XX AC AAK25658;

XX DT 21-MAY-1999 (first entry)

XX DE Human endogenous retrovirus W clone c1.P122.

XX KW Clone; human endogenous retrovirus; genome; autoimmune disease;

XX multiple sclerosis; rheumatoid polyarthritis; insulin-dependent diabetes;

XX disseminated lupus erythematosus; pregnancy; chromosomal marker; ss.

XX OS Human endogenous retrovirus.

XX WO9902696-A1.

XX 21-JAN-1999.

XX 06-JUL-1998; 98WO-FR01442.

XX 07-JUL-1997; 97FR-0008815.

XX PR

XX XX

XX XX

XX XX

PA (INNR) BIO MERIEUX.
XX Beseme F, Blond JL, Bouton O, Mallet F, Mandrand B;
XX WPI: 1999-120897/10.
XX
XX New nucleic acid sequences from human endogenous retrovirus-W -
PT expressed exclusively in placenta and useful in diagnosis and
PT therapy of autoimmune disease, and abnormal or failed pregnancy
XX
XX Claim 1; Page 54-56; 106pp; French.
XX
XX This sequence represents clone c1.P122 of the human endogenous retrovirus
CC (HERV) W genome. The nucleic acids, their fragments or peptides encoded
CC by them are markers of autoimmune disease (e.g. multiple sclerosis,
CC rheumatoid polyarthritis, disseminated lupus erythematosus, insulin-
CC dependent diabetes and related pathologies) and of abnormal or
CC unsuccessful pregnancy and can be used as chromosomal markers for
CC susceptibility to these conditions, or proximity markers of genes
CC associated with this susceptibility.
XX
SQ Sequence 2006 BP; 625 A; 532 C; 373 G; 475 T; 1 other;

Query Match 98.4%; Score 197.8; DB 20; Length 2006;
Best Local Similarity 99.0%; Pred. No. 1.4e-53;
Matches 199; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CAACCTAGTGGAGACATCACCCTCCTTAGCCAAATATCAACAAGTCTTAAACATTACA 60
DB 1042 CAACCTAGTGGAGACATCACCCTCCTTAGCCAAATATCAACAAGTCTTAAACATTACA 1101
QY 61 AGAAGCTATTCCTCGAGAAGAGGAAAGAACTATTCACCCCTGTGACATGTAATTAGT 120
DB 1102 AGAAGCTATTCCTCGAGAAGAGGAAAGAACTATTCACCCCTGTGACATGTAATTAGT 1161
QY 121 CAAGTCCCTTCCTCTTAATTCCTCATATCATATCCTGGGAAGACCTACCCAGT 180
DB 1162 CAAGTCCCTTCCTCTTAATTCCTCATATCATATCCTGGGAAGACCTACCCAGT 1221
QY 181 CATTTATCTACCCCACTGC 201
DB 1222 CATTTATCTACCCCACTGC 1242

RESULT 11
AAS59208
ID AAS59208 standard; DNA: 2006 BP.
XX
AC AAS59208;
XX
DT 07-NOV-2000 (first entry)
XX
DE 3' pol gene and 5' env gene sequences of HERV-W from human genome.
XX
KW Autoimmune disease; retrovirus; human endogenous retrovirus W; HERV-W;
KW gag gene; pregnancy; multiple sclerosis; T cell proliferation; ss.
XX
OS Homo sapiens.
XX
PN WO200043521-A2.
XX
PD 27-JUL-2000.
XX
PE 21-JAN-2000; 2000WO-FR00144.
XX
PR 21-JAN-1999; 99FR-0000888.
XX
PA (INNR) BIO MERIEUX.
XX
PI Paranhos-Baccala G, Mallet F, Voisset C;
XX
DR WPI: 2000-499229/44.
XX

PT New nucleic acid from human endogenous retrovirus, useful e.g. for
PT diagnosis of autoimmune disease and complications of pregnancy,
PT contains at least part of the gag gene -
XX
XX Disclosure; Page 44-45; 53pp; French.
XX
XX The present sequence represents an endogenetic retroviral nucleic acid
CC fragment, which is associated with an autoimmune disease, and is
CC integrated into the human genome. The fragment is originally derived
CC from a novel retrovirus, human endogenous retrovirus W (HERV-W). The
CC HERV-W retrovirus is associated with autoimmune disease, failure of
CC pregnancy or disorders of pregnancy. The nucleic acid fragment, or
CC proteins derived from it, are useful for diagnosis of autoimmune
CC disease (specifically multiple sclerosis) and for monitoring pregnancy.
CC The nucleic acid fragments may also be used for in situ labelling of
CC isolated chromosomes, while the transcription product can be used to
CC study or monitor T cell proliferation in vitro.
XX
SQ Sequence 2006 BP; 625 A; 532 C; 373 G; 475 T; 1 other;

Query Match 98.4%; Score 197.8; DB 21; Length 2006;
Best Local Similarity 99.0%; Pred. No. 1.4e-53;
Matches 199; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CAACCTAGTGGAGACATCACCCTCCTTAGCCAAATATCAACAAGTCTTAAACATTACA 60
DB 1042 CAACCTAGTGGAGACATCACCCTCCTTAGCCAAATATCAACAAGTCTTAAACATTACA 1101
QY 61 AGAAGCTATTCCTCGAGAAGAGGAAAGAACTATTCACCCCTGTGACATGTAATTAGT 120
DB 1102 AGAAGCTATTCCTCGAGAAGAGGAAAGAACTATTCACCCCTGTGACATGTAATTAGT 1161
QY 121 CAAGTCCCTTCCTCTTAATTCCTCATATCATATCCTGGGAAGACCTACCCAGT 180
DB 1162 CAAGTCCCTTCCTCTTAATTCCTCATATCATATCCTGGGAAGACCTACCCAGT 1221
QY 181 CATTTATCTACCCCACTGC 201
DB 1222 CATTTATCTACCCCACTGC 1242

RESULT 12
AAS84200
ID AAS84200 standard; CDNA: 1148 BP.
XX
AC AAS84200;
XX
DT 13-FEB-2002 (first entry)
XX
DE DNA encoding novel human diagnostic protein #20004.
XX
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX
OS Homo sapiens.
XX
PN WO200175067-A2.
XX
PD 11-OCT-2001.
XX
PE 30-MAR-2001; 2001WO-US08631.
XX
PR 31-MAR-2000; 2000US-0540217.
XX
PR 23-AUG-2000; 2000US-0649167.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Drmanac RT, Liu C, Tang YT;
XX
DR WPI: 2001-639362/73.
XX
DR P-PSDB; ABG20013.
XX
XX New isolated polynucleotide and encoded polypeptides, useful in

PT diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity

PS Claim 1: SEQ ID NO 20004; 103pp; English.

XX The invention relates to isolated polynucleotide (I) and
 CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
 CC polymerase chain reaction (PCR) primers, oligomers and for chromosome
 CC and gene mapping, and in recombinant production of (II). The
 CC polynucleotides are also used in diagnostics as expressed sequence tags
 CC for identifying expressed genes. (I) is useful in gene therapy techniques
 CC to restore normal activity of (II) or to treat disease states involving
 CC (II). (II) is useful for generating antibodies against it, detecting or
 CC quantitating a polypeptide in tissue, as molecular weight markers and as
 CC a food supplement. (II) and its binding partners are useful in medical
 CC imaging of sites expressing (II). (I) and (II) are useful for treating
 CC disorders involving aberrant protein expression or biological activity.
 CC The polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data or products dependent on DNA and
 CC amino acid sequences. AAS64197-AAS94514 represent novel human
 CC diagnostic coding sequences of the invention.
 CC Note: The sequence data for this patent did not appear in the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 CC XX

SO Sequence 1148 BP; 368 A; 286 C; 242 G; 252 T; 0 other;

Query Match 92.0%; Score 185; DB 23; Length 1148;
 Best Local Similarity 95.0%; Pred. No. 1,5e-49;
 Matches 191; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 1 CAACCTAGTTGCGACATCATCCTCTTAGCCAAATATCAACAAGTTCTTAAACATTACA 60
 DB 797 CAACCTAGTTGCGACATCATCCTCTTAGCCAAATATCAACAAGTTCTTAAACATTACA 856
 QY 61 AGAACCCTATCCCTGAGAGAGGAGAAAGAACTATTCCACCTCTGTGACATGTTATTAGT 120
 DB 857 AGAGGCTGTGCCCCGAGAGAGGAGAAAGAAATTTCCACCTGTGTCTATGTTATTAGT 916
 QY 121 CAAGTCCCTTCCCTTAATTTCCCATCCCTAGATACATCTGTGGAGAGACCTTACCAGT 180
 DB 917 CAAGTCCCTTCCCTTAATTTCCCATCCCTAGATACATCTGTGGAGAGACCTTACCAGT 976
 QY 181 CATTATCTACCCCAACTGC 201
 DB 977 CATTATCTATCCCAACTGC 997

RESULT 13

AAK25668
 ID AAK25668 standard; cDNA to mRNA; 2575 Bf.

AC AAK25668;

DT 21-MAY-1999 (first entry)

DE Human endogenous retrovirus W clone cl.7A21.

XX Clone; human endogenous retrovirus; genome; autoimmune disease;
 KW multiple sclerosis; rheumatoid polyarthriti; insulin-dependent diabetes;
 KW disseminated lupus erythematosus; pregnancy; chromosomal marker; ss.

OS Human endogenous retrovirus.

PN WO9902696-A1.

PD 21-JAN-1999.

PF 06-JUL-1998; 98WO-FR01442.

XX

PR 07-JUL-1997; 97FR-0008815.

PA (INMR) BIO MERIEUX.

XX Beseme F, Blond JL, Bouton O, Mallet F, Mandrand B;

PI WPI: 1999-120897/10.

DR New nucleic acid sequences from human endogenous retrovirus-W -

PT expressed exclusively in placenta and useful in diagnosis and

PI therapy of autoimmune disease, and abnormal or failed pregnancy

PS Claim 1: Page 80-83; 106pp; French.

XX This sequence represents clone cl.7A21 of the human endogenous retrovirus
 CC (HERV) W genome. The nucleic acids, their fragments or peptides encoded
 CC by them are markers of autoimmune disease (e.g. multiple sclerosis,
 CC rheumatoid polyarthriti, disseminated lupus erythematosus, insulin-
 CC dependent diabetes and related pathologies) and of abnormal or
 CC unsuccessful pregnancy and can be used as chromosomal markers for
 CC susceptibility to these conditions, or proximity markers of genes
 CC associated with this susceptibility.

SO Sequence 2575 BP; 817 A; 608 C; 546 G; 602 T; 2 other;

Query Match 89.7%; Score 180.2; DB 20; Length 2575;
 Best Local Similarity 93.5%; Pred. No. 7.1e-48;
 Matches 188; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

QY 1 CAACCTAGTTGCGACATCATCCTCTTAGCCAAATATCAACAAGTTCTTAAACATTACA 60
 DB 2285 CAACCTAGTTGCGACATCATCCTCTTAGCCAAATATCAACAAGTTCTTAAACATTACA 2344
 QY 61 AGAACCCTATCCCTGAGAGAGGAGAAAGAACTATTCCACCTCTGTGACATGTTATTAGT 120
 DB 2345 AGAGGCTGTGCCCCGAGAGAGGAGAAAGAACTATTCCACCTGTGTGACATGTTATTAGT 2404
 QY 121 CAAGTCCCTTCCCTTAATTTCCCATCCCTAGATACATCTGTGGAGAGACCTTACCAGT 180
 DB 2405 CAAGTCCCTTCCCTTAATTTCCCATCCCTAGATACATCTGTGGAGAGACCTTACCAGT 2464
 QY 181 CATTATCTACCCCAACTGC 201
 DB 2465 CATTATCTATCCCAACTGC 2485

RESULT 14

AAK25663
 ID AAK25663 standard; cDNA to mRNA; 3372 Bf.

AC AAK25663;

DT 21-MAY-1999 (first entry)

DE Human endogenous retrovirus W clone cl.P15T.

XX Clone; human endogenous retrovirus; genome; autoimmune disease;
 KW multiple sclerosis; rheumatoid polyarthriti; insulin-dependent diabetes;
 KW disseminated lupus erythematosus; pregnancy; chromosomal marker; ss.

OS Human endogenous retrovirus.

PN WO9902696-A1.

PD 21-JAN-1999.

PF 06-JUL-1998; 98WO-FR01442.

PR 07-JUL-1997; 97FR-0008815.

PA (INMR) BIO MERIEUX.

PI Beseme F, Blond JL, Bouton O, Mallet F, Mandrand B;

XX

XX DR WPI; 1999-120897/10.
XX PT New nucleic acid sequences from human endogenous retrovirus-W -
XX PT expressed exclusively in placenta and useful in diagnosis and
XX PT therapy of autoimmune disease, and abnormal or failed pregnancy
XX PS Claim 1; Page 64-67; 106pp; French.
XX This sequence represents clone cl.P15F of the human endogenous retrovirus
CC (HERV) W genome. The nucleic acids, their fragments or peptides encoded
CC by them are markers of autoimmune disease (e.g. multiple sclerosis,
CC rheumatoid polyarthritis, disseminated lupus erythematosus, insulin-
CC dependent diabetes and related pathologies) and of abnormal or
CC unsuccessful pregnancy and can be used as chromosomal markers for
CC susceptibility to these conditions, or proximity markers of genes
CC associated with this susceptibility.
XX SQ Sequence 3372 BP; 1047 A; 835 C; 711 G; 779 T; 0 other;
XX
XX Query Match 89.1%; Score 179; DB 20; Length 3372;
XX Best Local Similarity 96.5%; Pred. No. 1.9e-47;
XX Matches 194; Conservative 0; Mismatches 5; Indels 2; Gaps 1;
XX
XX QY 1 CAACTTAGTGCAGACATCACTCTTAGCCAAATATCAACAAGTCTTAAACATTACA 60
XX DB 2431 CAACTTAGTGCAGACATCACTCTTAGCCAAATATCAACAAGTCTTAAACATTACA 2490
XX
XX QY 61 AGGAACCTATCCCTGAGAGAAGGAGAAAGAACTATTCACCCCTTGATGATTTAGT 120
XX DB 2491 AGGAACCTATCCCTGAGAGAAGG--AAAAGAAATATTCACCCCAAGTGTATTAGT 2548
XX
XX QY 121 CAACTCCCTTCCCTCTAATTCCTCCATCCCTAGATACATCCTGGGAAGACCTACCCAGT 180
XX DB 2549 CAACTCCCTTCCCTCTAATTCCTCCATCCCTAGATACATCCTGGGAAGACCTACCCAGT 2608
XX
XX QY 181 CATTATATCTACCCCAACTGC 201
XX DB 2609 CATTATATCTACCCCAACTGC 2629
XX
XX RESULT 15
XX ID AAA59213 standard; DNA; 3372 BP.
XX AC AAA59213;
XX
XX DT 07-NOV-2000 (first entry)
XX
XX DE Partial pol gene and U3-R region sequences of HERV-W from human genome.
XX
XX KM Autoimmune disease; retrovirus; human endogenous retrovirus W; HERV-W;
XX KM gag gene; pregnancy; multiple sclerosis; T cell proliferation; ss.
XX
XX OS Homo sapiens.
XX
XX PN WC200043521-A2.
XX
XX PD 27-JUL-2000.
XX
XX PF 21-JAN-2000; 2000WO-FR00144.
XX
XX PR 21-JAN-1999; 99FR-0000888.
XX
XX PA (INMR) BIO MERIEUX.
XX
XX PI Paranhos-Baccala G, Mallet F, Voisset C;
XX
XX DR WPI; 2000-499229/44.
XX
XX PT New nucleic acid from human endogenous retrovirus, useful e.g. for
XX PT diagnosis of autoimmune disease and complications of pregnancy,
XX PT contains at least part of the gag gene

XX PS Disclosure; Page 47-48; 53pp; French.
XX
XX CC The present sequence represents an endogenous retroviral nucleic acid
XX CC fragment, which is associated with an autoimmune disease, and is
XX CC integrated into the human genome. The fragment is originally derived
XX CC from a novel retrovirus, human endogenous retrovirus W (HERV-W). The
XX CC HERV-W retrovirus is associated with autoimmune disease, failure of
XX CC pregnancy or disorders of pregnancy. The nucleic acid fragment, or
XX CC proteins derived from it, are useful for diagnosis of autoimmune
XX CC disease (specifically multiple sclerosis) and for monitoring pregnancy.
XX CC The nucleic acid fragments may also be used for in situ labelling of
XX CC isolated chromosomes, while the transcription product can be used to
XX CC study or monitor T cell proliferation in vitro.
XX
XX SQ Sequence 3372 BP; 1047 A; 835 C; 711 G; 779 T; 0 other;
XX
XX Query Match 89.1%; Score 179; DB 21; Length 3372;
XX Best Local Similarity 96.5%; Pred. No. 1.9e-47;
XX Matches 194; Conservative 0; Mismatches 5; Indels 2; Gaps 1;
XX
XX QY 1 CAACTTAGTGCAGACATCACTCTTAGCCAAATATCAACAAGTCTTAAACATTACA 60
XX DB 2431 CAACTTAGTGCAGACATCACTCTTAGCCAAATATCAACAAGTCTTAAACATTACA 2490
XX
XX QY 61 AGGAACCTATCCCTGAGAGAAGGAGAAAGAACTATTCACCCCTTGATGATTTAGT 120
XX DB 2491 AGGAACCTATCCCTGAGAGAAGG--AAAAGAAATATTCACCCCAAGTGTATTAGT 2548
XX
XX QY 121 CAACTCCCTTCCCTCTAATTCCTCCATCCCTAGATACATCCTGGGAAGACCTACCCAGT 180
XX DB 2549 CAACTCCCTTCCCTCTAATTCCTCCATCCCTAGATACATCCTGGGAAGACCTACCCAGT 2608
XX
XX QY 181 CATTATATCTACCCCAACTGC 201
XX DB 2609 CATTATATCTACCCCAACTGC 2629
XX

Search completed: May 2, 2003, 14:51:19
Job time : 184.333 secs

GenCore version 5.1.5
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 2, 2003, 14:19:30 ; Search time 1112 Seconds
(without alignments)
2927.422 Million cell updates/sec

Title: US-09-719-554-3_COPY_7500_7700

Perfect score: 201
Sequence: 1 caactagttgcagacatca.....atttattaccccaactgc 201

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

EST: *
1: em_estba: *
2: em_esthum: *
3: em_estin: *
4: em_estnu: *
5: em_estov: *
6: em_estpl: *
7: em_estro: *
8: em_hlc: *
9: gb_est1: *
10: gb_est2: *
11: gb_hlc: *
12: gb_est3: *
13: gb_est4: *
14: gb_est5: *
15: em_estfun: *
16: em_estom: *
17: gb_gss: *
18: em_gss_hum: *
19: em_gss_inv: *
20: em_gss_pln: *
21: em_gss_vrt: *
22: em_gss_fun: *
23: em_gss_mam: *
24: em_gss_mus: *
25: em_gss_other: *
26: em_gss_pro: *
27: em_gss_rod: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	201	100.0	362	13	BI049622 CM3-GN029
2	201	100.0	503	9	BG000981 I10-GN024
3	177	88.1	578	9	AA613238 no19f04.s
4	153	76.1	414	14	HI2439 yj11n10.r1
5	144.8	72.0	171	12	BG013341 CM4-GN036
6	101.8	50.6	605	17	AQ240663 CIT-HSP-2

Result No.	Score	Query Match	Length	ID	Description
7	97.6	48.6	496	10	BE208551 ba09a03.Y
8	97	48.3	618	17	B59957 CIT-HSP-347
9	95.2	47.4	516	17	AQ428067 CITR1-E1
10	91	45.3	536	10	AM074844 xa08c04.x
11	89	44.3	672	17	B68542 CIT-HSP-205
12	86.4	43.0	138	12	BG013345 CM4-GN036
13	84.6	41.1	578	12	B54046 CIT-HSP-201
14	82.6	41.1	563	17	AQ427133 CITR1-E1
15	80	39.8	477	17	AQ309705 CIT-HSP-2
16	78.8	39.2	330	14	N76991 yv50a07.r1
17	70.6	35.1	483	17	AQ609732 HS-5076.A
18	69.2	34.4	408	13	BI034898 QV2-NN200
19	67.4	33.5	347	14	BQ024998 UT-1-BB1P
20	62	30.8	403	17	AQ632527 RPT1-11-4
21	61.4	30.5	426	13	BI033854 QV2-NN200
22	60	29.9	723	13	BI772113 603059196
23	57	28.4	448	17	BI7253 344P9.TV.C
24	55.8	27.8	574	17	AQ634644 RPT1-11-4
25	55.8	27.8	637	17	AQ378897 RPT1-11-16
26	48.6	24.2	422	17	AQ228090 HS-2014.B
27	48	23.9	363	17	AQ061656 CIT-HSP-2
28	48	23.9	390	17	AQ059215 CIT-HSP-2
29	47	23.4	751	13	BM014497 603640281
30	46	22.9	668	12	BE730884 601570001
31	43.8	21.8	451	17	AQ184628 HS-2212.A
32	40.4	20.1	290	14	BQ022458 UT-1-BB1P
33	39.2	19.5	576	17	B52708 CIT-HSP-200
34	39	19.4	561	17	AQ487402 RPT1-11-2
35	39	18.9	614	17	B75173 RPT1-11-2
36	38	18.4	939	17	AL059400 Drosophila
37	36.2	18.0	700	17	AQ262910 CITR1-E1
38	35	17.4	609	17	A2796232 2M051D22
39	35	17.4	710	13	BG914193 602809130
40	35	17.4	1735	12	BE790606 601582961
41	34.8	17.3	546	10	BE684788 186498.MA
42	34.6	17.2	513	9	AA498640 vhl3f03.r
43	34.6	17.2	629	17	AG125475 Pan trogl
44	34.4	17.1	438	17	B99018 CIT-HSP-228
45	34	16.9	945	17	AL251887 Tetradon

ALIGNMENTS

RESULT 1
LOCUS BI049622 362 bp mRNA linear EST 15-JUN-2001
DEFINITION CM3-GN0297-020101-591-g08 GN0297 Homo sapiens cDNA, mRNA sequence.
ACCESSION BI049622
VERSION BI049622.1 GI:14457152
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 362)
AUTHORS Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Brites,M.R., Nagai,M.A., da Silva,W.U., Zago,M.A., Bordin,S., Costa,F.F., Goldman,G.H., Carvalho,A.F., Matsukuma,A., Bata,G.S., Simpson,D.H., Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J., and Simpson,A.J.
TITLE Shotgun sequencing of the human transcriptome with ORF expressed
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
MEDLINE 2020263
COMMENT Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001

Email: asimpson@ludwig.org.br
 This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL:
 (http://www.ludwig.org.br/scripts/gethtml2.pl?tbl=CM3&cl=CM3-GN0297-020101-591-d08&cl=2001-01-02&cl=1)
 Seq primer: puc 18 forward
 High quality sequence start: 3
 High quality sequence stop: 362.
 Location/Qualifiers
 source
 1..362
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_lib="GN0297"
 /dev_stage="Adult"
 /note="Organ: placenta.normal; Vector: puc18; Site_1: SmaI; Site_2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

BASE COUNT 111 a 100 c 64 g 87 t
 ORIGIN

Query Match 100.0%; Score 201; DB 13; Length 362;
 Best Local Similarity 100.0%; Pred. No. 1.5e-51;
 Matches 201; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 CAACCTAGTGCAGACATCAGCTCCTTAGCCAAATATCAACAGTTCTTAAACATTACA 60
 Db 140 CAACCTAGTGCAGACATCAGCTCCTTAGCCAAATATCAACAGTTCTTAAACATTACA 199
 QY 61 AGAACCTATCCCTGGAGAAAGGAAAGAAAGAACTATCCACCCTTGACATGGATTAGT 120
 Db 200 AGAACCTATCCCTGGAGAAAGGAAAGAAAGAACTATCCACCCTTGACATGGATTAGT 259
 QY 121 CAAGTCCTTCCTCTATATTCCTCCTAGATATATCTCTGGAGAGACCTTACCACT 180
 Db 260 CAAGTCCTTCCTCTATATTCCTCCTAGATATATCTCTGGAGAGACCTTACCACT 319
 QY 181 CATTTTATCTACCCCAACTGC 201
 Db 320 CATTTTATCTACCCCAACTGC 340

RESULT 2
 BG000981 503 bp mRNA linear EST 24-JAN-2001
 LOCUS IL0-GN0246-131100-501-c06 GN0246 Homo sapiens cDNA, mRNA sequence.
 DEFINITION BG000981 GI:12438850
 ACCESSION
 VERSION
 KEYWORDS
 EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 1 (bases 1 to 503)
 Dias Neto, E., Garcia Correa, R., Verjowski-Almeida, S., Briones, M.R., Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F., Goldman, G.H., Carvalho, A.F., Matsukuma, A., Bata, G.S., Simpson, D.H., Brunstein, A., deoliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare, M.J., Soares, F., Brentani, R.R., Neils, L.F., de Souza, S.J. and Simpson, A.J.
 Shotgun sequencing of the human transcriptome with ORF expressed sequence tags
 Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
 JOURNAL
 MEDLINE
 COMMENT
 Contact: Simpson A.J.G.
 Laboratory of Cancer Genetics
 Ludwig Institute for Cancer Research
 Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil
 Tel: +55-11-2704922

Fax: +55-11-2707001
 Email: asimpson@ludwig.org.br
 This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL:
 (http://www.ludwig.org.br/scripts/gethtml2.pl?tbl=IL0&cl=IL0-GN0246-131100-501-c06&cl=2000-11-13&cl=1)
 Seq primer: puc 18 forward
 High quality sequence stop: 503.
 Location/Qualifiers
 source
 1..503
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_lib="GN0246"
 /dev_stage="Adult"
 /note="Organ: placenta.normal; Vector: puc18; Site_1: SmaI; Site_2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

BASE COUNT 159 a 134 c 90 g 120 t
 ORIGIN

Query Match 100.0%; Score 201; DB 12; Length 503;
 Best Local Similarity 100.0%; Pred. No. 1.7e-51;
 Matches 201; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 CAACCTAGTGCAGACATCAGCTCCTTAGCCAAATATCAACAGTTCTTAAACATTACA 60
 Db 127 CAACCTAGTGCAGACATCAGCTCCTTAGCCAAATATCAACAGTTCTTAAACATTACA 186
 QY 61 AGAACCTATCCCTGGAGAAAGGAAAGAAAGAACTATCCACCCTTGACATGGATTAGT 120
 Db 187 AGAACCTATCCCTGGAGAAAGGAAAGAAAGAACTATCCACCCTTGACATGGATTAGT 246
 QY 121 CAAGTCCTTCCTCTATATTCCTCCTAGATATATCTCTGGAGAGACCTTACCACT 180
 Db 247 CAAGTCCTTCCTCTATATTCCTCCTAGATATATCTCTGGAGAGACCTTACCACT 306
 QY 181 CATTTTATCTACCCCAACTGC 201
 Db 307 CATTTTATCTACCCCAACTGC 327

RESULT 3
 AA613238 578 bp mRNA linear EST 03-OCT-1997
 LOCUS n019f04.s1 NCI-CGAP_Phel Homo sapiens cDNA clone IMAGE:1101151 3'
 DEFINITION similar to SW:POL_SMGAV P03359 POL POLYPROTEIN ;, mRNA sequence.
 ACCESSION AA613238
 VERSION
 KEYWORDS
 EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 1 (bases 1 to 578)
 NCI-CGAP http://www.ncbi.nlm.nih.gov/ncigap
 National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
 Unpublished (1997)
 Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: W. Marston Linehan, M.D., Michael R. Emmert-Buck, M.D., Ph.D.
 CDNA Library Preparation: Stratagene, Inc.
 CDNA Library Arrayed by: Greg Lennon, Ph.D.
 DNA Sequencing by: Washington University Genome Sequencing Center
 Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
 www-bio.llnl.gov/bbrp/image/image.html
 Seq primer: -40m13 fwd. ET from Amersham

High quality sequence stop: 442.

FEATURES
Location/Qualifiers

1..578

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="IMAGE:1101151"

/tissue="pneumocystoma"

/lab_host="SOLR (kanamycin resistant)"

/note="Vector: Bluescript SK-; Site: EcoRI; Site 2: XhoI"

/cloned_unidirectionally. Primer: Oligo dT.

Pneumocystoma, 5' adaptor sequence: 5' GATTCGGCAGCAG

3' 3' adaptor sequence: 5' CTCGAGCTTTTCTTTTCTTTT 3'

Average insert size: 1.3 Kb."

BASE COUNT

128 a 116 c 160 g 174 t

ORIGIN

Query Match

Best Local Similarity 92.5%; Score 177; DB 9; Length 578;

Matches 186; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

OY 1 CACTTAGTTCAGACATCACCCTCTTACCAATATACAGTTCTTAAACATTCACA 60

DB 237 CAACCTAGTTCAGATATACCTCTTACCAATATACAGTTCTTAAACATTCACA 178

OY 61 AGAAGCTATCCCTGAGAGAGGAGAAAGAACTATTCACCTGTGACATGTTAGT 120

DB 177 AGAAGCTATCCCTGAGAGAGGAGAAAGAACTATTCACCTGTGACATGTTAGT 118

OY 121 CAAGTCCCTTCCCTTAATTCCTTACCTGATACATCTCTGGAAGAGCCCTACCT 180

DB 117 CAAGTCCCTTCCCTTCTGATTCCTTACCTGATACATCTCTGGAAGAGCCCTACCT 58

OY 181 CATTTATCTACCCCACTGC 201

DB 57 CATTTATCTACCCCACTGC 37

RESULT 4

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

H12439 414 bp mRNA linear EST 27-JUN-1995

Y111h10.r1 Soares placenta ND2HP Homo sapiens cDNA clone

IMAGE:148483 5', mRNA sequence.

H12439

H12439.1 GI:877259

EST.

human.

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.

1 (bases 1 to 414)

Hillier, L., Clark, N., Dubuque, T., Elliston, K., Hawkins, M., Holman

, M., Hultman, M., Kucaba, T., Le, M., Lennon, G., Matris, M., Parsons, J.,

Rifkin, L., Rohlfing, T., Soares, M., Tan, F., Trevaskis, E., Waterston

, R., Williamson, A., Wohlmann, P. and Wilson, R.

The WashU-Merck EST Project

Unpublished (1995)

Contact: Wilson RK

Washington University School of Medicine

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108

Tel: 314 286 1800

Fax: 314 286 1810

Email: est@watson.wustl.edu

Insert size: 678

High quality sequence stops: 257

Source: IMAGE Consortium, LNL

This clone is available royalty-free through LNL; contact the

IMAGE Consortium (info@image.lnl.gov) for further information.

FEATURES

Source

Location/Qualifiers

1..414

/organism="Homo sapiens"

/db_xref="GDB:560252"

/db_xref="taxon:9606"

/clone="IMAGE:148483"

/clone_id="Soares placenta ND2HP"

/sex="Female"

/dev_stage="placenta obtained at birth (full term)"

/lab_host="DH10B (ampicillin resistant)"

/note="Organ: placenta; Vector: pT73D (Pharmacia) with a

modified polylinker; Site: Not I; Site 2: Eco RI; 1st

strand cDNA was primed with a Not I - oligo(dT) primer [5'

ACTGGAAGAAATTCGCGCGCAGAGAAATTTTCTTTTCTTTT 3'],

double-stranded cDNA was ligated to Eco RI adaptors

(Pharmacia), digested with Not I and cloned into the Not I

and Eco RI sites of the modified pT73 vector. Library

went through one round of normalization. Library

constructed by Bento Soares and M.Fatima Bonaldo.

Query Match

Best Local Similarity 99.4%; Score 153; DB 14; Length 414;

Matches 164; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

OY 1 CACTTAGTTCAGACATCACCCTCTTACCAATATACAGTTCTTAAACATTCACA 60

DB 149 CACTTAGTTCAGACATCACCCTCTTACCAATATACAGTTCTTAAACATTCACA 208

OY 61 AGAAGCTATCCCTGAGAGAGGAGAAAGAACTATTCACCTGTGACATGTTAGT 120

DB 209 AGAAGCTATCCCTGAGAGAGGAGAAAGAACTATTCACCTGTGACATGTTAGT 268

OY 121 CAAGTCCCTTCCCTTAATTCCTTACCTGATACATCTCTGGAAGAGCCCTACCT 164

DB 269 CAAGTCCCTTCCCTTCTGATTCCTTACCTGATACATCTCTGGAAGAGCCCTACCT 313

RESULT 5

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

COMMENT

BG013341 171 bp mRNA linear EST 24-JAN-2001

CM4-GN0361-161200-628-a02 GN0361 Homo sapiens cDNA, mRNA sequence.

BG013341

BG013341.1 GI:12463434

EST.

human.

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.

1 (bases 1 to 171)

Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R.,

Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F.,

Goldman, G.H., Carvalho, A.F., Matsukuma, A., Bata, G.S., Simpson, D.H.,

Brunstein, A., de Oliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare

, M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and

Simpson, A.J.

Shotgun sequencing of the human transcriptome with ORF expressed

sequence tags

Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)

Contact: Simpson A.J.G.

Laboratory of Cancer Genetics

Ludwig Institute for Cancer Research

Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,

Brazil

Tel: +55-11-2704922

Fax: +55-11-2707001

Email: asimpson@ludwig.org.br

This sequence was derived from the FAPESP/LICR Human Cancer Genome

Project. This entry can be seen in the following URL

(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=CM4&tl2=CM4-GN0361-

161200-628-a02&tl3=2000-12-16&tl4=1)

Seg primer: puc 18 forward

High quality sequence start: 13

BASE COUNT 111 a 101 c 126 g 157 t 1 others
(Stratagene) and Superscript II RT (Life Technologies)."

Query Match 48.6%; Score 97.6; DB 10; Length 496;
Best Local Similarity 67.7%; Pred. No. 1.6e-19;
Matches 136; Conservative 0; Mismatches 65; Indels 0; Gaps 0;

QY 1 CAATTAGTGCAGACATCACTCTTACGCAATATACACAGTTCTTAAACATTACA 60
|||||
DB 371 CAACTTGTGCAAAATATTAATCTTGTGGCAAAAANCAAGAAACCTTAAACCTAC 312
AGGACATATCCCTGAGAGAGAGAGAAAGAACTATTCACCCCTGTGACATGATTAGT 120
DB 311 TGAAGATGTCACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 252
QY 121 CAACTCCCTTCCCTCTTAATTCCTCCATCCCTAGATACATCTGGAGAGACCTACCCAGT 180
DB 251 CAAATCTCTCCCTCTACCTCCATCAATGATTCTTGTGGAGAGAGCAAACTCGGT 192
QY 181 CATTTTATCTACCCCACTGC 201
DB 191 AATCCTCTACCCCACTGC 171

RESULT 8

B59957 618 bp DNA linear GSS 21-JUN-1998
LOCUS CIT-HSP-347118.TVC CIT-HSP Homo sapiens genomic clone 347118, DNA
DEFINITION sequence.

ACCESSION B59957
VERSION B59957.1 GI:2614675
KEYWORDS GSS.
SOURCE human.
ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
AUTHORS 1 (bases 1 to 618)
Adams,M.D., Rounsley,S.D., Field,C.E., Baas,S., Linher,K., Golden
,K., Berry,K., Grainger,D., Sun,E., Wilde,C., Shizuya,H., Simon,M.
and Venter,J.C.
TITLE Use of a random BAC End Sequence Database for Sequence-Ready Map
Building

JOURNAL Unpublished (1997)
COMMENT Other-GSS: 347118.TP 347118.TV
Contact: Mark Adams
Department of Eukaryotic Genomics
The Institute for Genomic Research
7912 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: mdadams@tigr.org

Clones are available from Research Genetics (info@resgen.com). BAC
end search page:
http://www.tigr.org/tldb/hungen/bac_end_search/bac_end_search.html
Seq primer: T7
Class: BAC ends.

FEATURES

Source Location/Qualifiers
1..618
/organism="Homo sapiens"
/db_xref="GDB:5364118"
/db_xref="taxon:9606"
/clone_id="347118"
/clone_lib="CIT-HSP"
/sex="Male"
/cell_type="Sperm"
/note="Vector: pBelobAC1; site_1: HindIII; site_2:
HindIII"

BASE COUNT 192 a 153 c 120 g 153 t
ORIGIN

Query Match 48.3%; Score 97; DB 17; Length 618;
Best Local Similarity 67.7%; Pred. No. 2.7e-19;

Matches 136; Conservative 0; Mismatches 65; Indels 0; Gaps 0;

QY 1 CAATTAGTGCAGACATCACTCTTACGCAATATACACAGTTCTTAAACATTACA 60
|||||
DB 300 CAACTTGTGCAAAATATTAATCTTGTGGCAAAAATATCAAGAAACCTTAAACCTAC 359
QY 61 AGGAACCTATCCCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 120
DB 360 TGAAGATGTCACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 419
QY 121 CAACTCCCTTCCCTCTTAATTCCTCCATCCCTAGATACATCTGGAGAGACCTACCCAGT 180
DB 420 CAAATCTCTCCCTCTACCTCCATCAATGATTCTTGTGGAGAGAGACATCTCGT 479
QY 181 CATTTTATCTACCCCACTGC 201
DB 480 AATCCTCTACCTCCCACTGC 500

RESULT 9

A0428067 516 bp DNA linear GSS 24-MAR-1999
LOCUS CITBI-EI-2566F23.TR CITBI-EI Homo sapiens genomic clone 2566F23,
DEFINITION DNA sequence.

ACCESSION A0428067
VERSION A0428067.1 GI:4500974
KEYWORDS GSS.
SOURCE human.
ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
AUTHORS Zhao,S., Adams,M.D., Nierman,W., Malek,J., Shizuya,H., Simon,M. and
Venter,J.C.
TITLE Use of BAC End Sequences from Caltech Libraries for Sequence-Ready
Map Building

JOURNAL Unpublished (1997)
COMMENT Other-GSS: CITBI-EI-2566F23.TR
Contact: Shaying Zhao, William Nierman, Mark Adams
Department of Eukaryotic Genomics
The Institute for Genomic Research
7912 Medical Center Dr., Rockville, MD 20850
Tel: 301 838 0200
Fax: 301 838 0208
Email: hbeet@tigr.org

Clones are available from Research Genetics (info@resgen.com). BAC
end search page:
http://www.tigr.org/tldb/hungen/bac_end_search/bac_end_search.html.
Seq primer: M13-21
Class: BAC ends.

FEATURES

Source Location/Qualifiers
1..516
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_id="2566F23"
/clone_lib="CITBI-EI"
/sex="Male"
/cell_type="Sperm"
/note="Vector: pBelobAC1; site_1: EcoRI; site_2: EcoRI;
Caltech Human BAC Library D"

Query Match 47.4%; Score 95.2; DB 17; Length 516;
Best Local Similarity 92.6%; Pred. No. 9e-19;
Matches 100; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 94 ATTCCACCTGTGTACATGATATAGTCAATCCCTTCCCTTAATTCCTCATCTCTAGA 153
DB 4 ATTCACCTGTGTACATGATATAGTCAATCCCTTCCCTTAATTCCTCATCTCTAGA 63
QY 154 TACATCTCTGTGTACATGATATAGTCAATCCCTTCCCTTAATTCCTCATCTCTAGA 201
|||||

DB 64 TACATACGTGGAGAGACCTACCCAGCTCATCTTATCTACCTGATGTC 111

RESULT 10

AM074844/c

LOCUS

AM074844

DEFINITION

AM074844

ACCESSION

AM074844

KEYWORDS

EST

SOURCE

human

ORGANISM

human sapiens

REFERENCE

1 (bases 1 to 536)

AUTHORS

NCI-CGAP

TITLE

NCI-CGAP

JOURNAL

Unpublished (1997)

COMMENT

Contact: Robert Strausberg, Ph.D.
Email: cgapbs-rt@mail.nih.gov
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Insert length: 831 Std Error: 0.00
Seq primer: -400p from g1bco
High quality sequence stop: 459.

FEATURES

1..536

source

Location/Qualifiers

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="IMAGE:2567718"

/clone_1lib="Soares_NFL_T_GBC_S1"

/lab_host="DH10B"

/note="Organ: pooled; vector: pT73D-Pac (Pharmacia) with
a modified polylinker; Site_1: Not I; Site_2: Eco RI;
Equal amounts of plasmid DNA from three normalized
libraries (fetal lung fibroblasts, testis NHT, and B-cell
NCI-CGAP (GCB1)) were mixed, and ss circles were made in
vitro. Following HAP purification, this DNA was used as
tracer in a subtractive hybridization reaction. The driver
was PCR-amplified cDNAs from pools of 5,000 clones made
from the same 3 libraries. The pools consisted of
I.M.A.G.E. clones 297430-302087, 682632-687239,
726408-728711, and 729036-731399. Subtraction by Bento
Soares and M. Fatima Bonaldo."

BASE COUNT

141 a 105 c 136 g 154 t

ORIGIN

Query Match

Best Local Similarity 66.7%; Score 91; DB 10; Length 536;

Matches 130; Conservative 0; Mismatches 65; Indels 0; Gaps 0;

1 CAACCTAGTGGAGACATCACTCTTACCCAAATTCACCAAGTCTTAAACATTACA 60

DB 405 CAATTTGGTCAATATATATCTTTGGCTAAATATCAATGAAACCTTAAACCCACC 346

DB 61 AGAAGCTATCCCTGAGAGAGAGAAAGAACTATTCACCCCTTGACATGCTATTAGT 120

DB 345 TGAAGATGTCAACAGAGAGAGAGAGAGAGTGTTCACACGAGAGATCTAGTGTGGT 286

DB 121 CAAGTCCCTCCCTCTAATTTCCCTCCTAGATATATCTGAGAGAGAGAGAGAGAGAG 180

DB 285 CAATATCTCTCCCTCTACCTCCCATCTATGATTTCTTTGTGGAGAGAGACATCACTCACT 226

DB 181 CATTATCTACCC 195

DB 225 AATCTCTCTGCTC 211

RESULT 11

B685442

LOCUS B68542 672 bp DNA linear GSS 21-JUN-1998

DEFINITION CIT-HSP-205314.TF CIT-HSP Homo sapiens genomic clone 205314, DNA

sequence.

ACCESSION B68542

VERSION B68542.1 GI:2667252

KEYWORDS GSS.

SOURCE human

ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

1 (bases 1 to 672)

Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

Adams,M.D., Rounsley,S.D., Field,C.E., Bass,S., Linher,K., Golden

K., Berry,K., Granger,D., Sun,E., Wible,C., Shizuya,H., Simon,M.

and Verity,J.C.

Use of a random BAC End Sequence Database for Sequence-Ready Map

Building

Unpublished (1997)

Contact: Mark Adams

Department of Eukaryotic Genomics

The Institute for Genomic Research

9712 Medical Center Dr., Rockville, MD 20850, USA

Tel: 301 838 0200

Fax: 301 838 0208

Email: mdadams@tigr.org

Clones are available from Research Genetics (info@resgen.com). BAC

end search page:

http://www.tigr.org/tdb/human/bac_end_search/bac_end_search.html

Seq primer: M13-21

Class: BAC ends.

1..672

/organism="Homo sapiens"

/db_xref="GDB:7057735"

/db_xref="taxon:9606"

/clone="205314"

/clone_1lib="CIT-HSP"

/sex="Male"

/cell_type="Sperm"

/note="Vector: pBelobAC11; Site_1: HindIII; Site_2: HindIII"

BASE COUNT 206 a 161 c 137 g 168 t

ORIGIN

Query Match

Best Local Similarity 65.2%; Score 89; DB 17; Length 672;

Matches 131; Conservative 0; Mismatches 70; Indels 0; Gaps 0;

1 CAACCTAGTGGAGACATCACTCTTACCCAAATTCACCAAGTCTTAAACATTACA 60

DB 457 CAATTTGGTCAATATATATCTTTGGCTAAATATCAATGAAACCTTAAACCCACC 516

DB 61 AGAAGCTATCCCTGAGAGAGAGAGAAAGAACTATTCACCCCTTGACATGCTATTAGT 120

DB 517 TGAAGATGTCAACAGAGAGAGAGAGAGAGTGTTCACACGAGAGATCTAGTGTGGT 576

DB 121 CAAGTCCCTCCCTCTAATTTCCCTCCTAGATATATCTGAGAGAGAGAGAGAGAGAG 180

DB 577 CAATATCTCTCCCTCTACCTCCCATCTATGATTTCTTTGTGGAGAGAGACATCACTCACT 636

DB 181 CATTATCTACCCCACTGC 201

DB 637 AATCTCTCTACCCAGACTGC 657

RESULT 12

B6013345

LOCUS B6013345 138 bp mRNA linear EST 24-JAN-2001

DEFINITION CM4-GN0361-161200-628-c08 GN0361 Homo sapiens cDNA, mRNA sequence.

ACCESSION B6013345

VERSION B6013345.1 GI:12463442

KEYWORDS EST.

SOURCE human

ORGANISM Homo sapiens

GenCore version 5.1.5
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 2, 2003, 14:52:21 ; Search time 39.333 Seconds
(without alignments)
1567.170 Million cell updates/sec

Title: US-09-719-554-3_COPY_7500_7700

Perfect score: 201
Sequence: 1 caactagttgcagacatca.....atttaccacccactgc 201

Scoring table: IDENTITY_NUC
Gapop 10.0 ; Gapext 1.0

Searched: 441362 seqs, 15338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : Issued_Patents_NA.*
1: /cgn2_6/ptodata/2/ina/5A_COMB.seq.*
2: /cgn2_6/ptodata/2/ina/5B_COMB.seq.*
3: /cgn2_6/ptodata/2/ina/6A_COMB.seq.*
4: /cgn2_6/ptodata/2/ina/6B_COMB.seq.*
5: /cgn2_6/ptodata/2/ina/PCTUS_COMB.seq.*
6: /cgn2_6/ptodata/2/ina/Backfile1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	201	100.0	2946	4	US-09-175-928-3
2	146.6	72.9	542	1	US-08-686-878A-48
3	146.6	72.9	542	1	US-08-721-489-2
4	29.2	14.5	50000	4	US-09-146-053-4
5	28.6	14.2	1564	4	US-09-071-035-63
6	28.6	14.2	1641	4	US-09-071-035-61
7	28.4	14.1	735	1	US-08-798-000-3
8	28.4	14.1	1234	1	US-08-798-000-1
9	28.4	14.1	2431	3	US-08-985-526-35
10	28.4	14.1	3612	4	US-09-265-013-3
11	28.4	14.1	5406	1	US-07-813-593-3
12	28.4	14.1	5406	1	US-07-977-451-5
13	28.4	14.1	5406	1	US-07-946-507-3
14	28.4	14.1	5406	1	US-08-252-517-5
15	28.4	14.1	5406	1	US-07-906-397A-5
16	28.4	14.1	5406	1	US-08-601-891-5
17	28.4	14.1	5406	2	US-09-021-324-5
18	28.4	14.1	5406	5	PCT-US92-02750-7
19	28.4	14.1	5406	5	PCT-US92-05401-5
20	28.4	14.1	5406	5	PCT-US92-09893-5
21	28.4	14.1	5470	2	US-08-443-861-1
22	28.4	14.1	5470	4	US-08-193-829B-1
23	28.4	14.1	6028	4	US-09-011-745-5
24	28.4	14.1	6061	4	US-09-011-745-6
25	28.4	14.1	6350	2	US-08-385-335A-8
26	28.4	14.1	7308	4	US-09-011-745-3
27	28.4	14.1	7308	4	US-09-011-745-4

28	28.4	14.1	7616	4	US-09-011-745-2	Sequence 2, Appli
29	28.4	14.1	8332	3	US-08-850-961-1	Sequence 1, Appli
30	28.4	14.1	8332	4	US-09-479-776-1	Sequence 1, Appli
31	28.4	14.1	8332	4	US-09-309-572-11	Sequence 11, Appli
32	28.4	14.1	8332	4	US-09-315-127-1	Sequence 1, Appli
33	28.4	14.1	8332	4	US-09-265-013-1	Sequence 1, Appli
34	28.4	14.1	8614	4	US-09-208-827-5	Sequence 5, Appli
35	28.4	14.1	10100	4	US-09-133-944-1	Sequence 1, Appli
36	28.4	14.1	10100	4	US-09-208-827-1	Sequence 1, Appli
37	28	13.9	568	4	US-09-404-879A-113	Sequence 113, App
38	27.8	13.8	11703	4	US-09-101-886B-3	Sequence 3, Appli
39	27.4	13.6	5057	2	US-08-365-486A-12	Sequence 12, Appli
40	27.4	13.6	5057	4	US-08-880-342-12	Sequence 12, Appli
41	27.4	13.6	5108	1	US-07-642-002-1	Sequence 1, Appli
42	27.4	13.6	10367	1	US-08-110-300A-9	Sequence 9, Appli
43	27.4	13.6	10367	2	US-08-886-642-9	Sequence 9, Appli
44	27.4	13.6	10367	5	PCT-US93-08041-9	Sequence 9, Appli
45	27.2	13.5	771	4	US-08-998-416-222	Sequence 222, App

ALIGNMENTS

```
RESULT 1
US-09-175-928-3
; Sequence 3, Application US/09175928A
; Patent No. 6312921
;
GENERAL INFORMATION:
; APPLICANT: Jacobs, Kenneth
; APPLICANT: McCoy, John M.
; APPLICANT: Lavallie, Edward R.
; APPLICANT: Collins-Racie, Lisa A.
; APPLICANT: Evans, Cheryl
; APPLICANT: Merberg, David
; APPLICANT: Treacy, Maurice
; APPLICANT: Mi, Sha
; APPLICANT: Genetics Institute, Inc.
; TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES ENCODING THEM
; FILE REFERENCE: 6006B.A172A
; CURRENT APPLICATION NUMBER: US/09/175,928A
; NUMBER OF SEQ ID NOS: 62
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 2946
; TYPE: DNA
; ORGANISM: Homo sapiens
;
Query Match          100.0%; Score 201; DB 4; Length 2946;
Best Local Similarity 100.0%; Pred. No. 6.2e-55;
Matches 201; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
;
QY 1 CAACCTAGTGGAGACATCCCTCTAGCCAAATATCAACAAGTCTTAAACATTACA 60
DB 549 CAACCTAGTGGAGACATCCCTCTAGCCAAATATCAACAAGTCTTAAACATTACA 608
;
QY 61 AGGAACCTATCCCTGAGAGAGAGAAAGAACTATTCACCCCTTGGAATGATTA 120
DB 609 AGGAACCTATCCCTGAGAGAGAGAAAGAAAGAACTATTCACCCCTTGGAATGATTA 668
;
QY 121 CAAGTCCCTCCCTCAATTCCTCCATCCCTAGATACATCTGGAGAGACCTACCCAGT 180
DB 669 CAAGTCCCTCCCTCAATTCCTCCATCCCTAGATACATCTGGAGAGACCTACCCAGT 728
;
QY 181 CATTTATCTACCCCACTGC 201
DB 729 CATTTATCTACCCCACTGC 749
;
RESULT 2
US-08-686-878A-48
; Sequence 48, Application US/08686878A
```

```

; Patent No. 5708157
; GENERAL INFORMATION:
; APPLICANT: Jacobs, Kenneth
; APPLICANT: McCoy, John
; APPLICANT: Lavallee, Edward
; APPLICANT: Racie, Lisa
; APPLICANT: Merberg, David
; APPLICANT: Treacy, Maurice
; APPLICANT: Evans, Cheryl
; APPLICANT: Spaulding, Vikki
; TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES
; NUMBER OF SEQUENCES: 71
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genetics Institute, Inc.
; STREET: 87 Cambridgepark Drive
; CITY: Cambridge
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02140
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/686,878A
; FILING DATE:
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Brown, Scott A.
; REGISTRATION NUMBER: 32,724
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 498-8224
; TELEFAX: (617) 876-5851
; INFORMATION FOR SEQ ID NO: 48:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 542 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
;
US-08-686-878A-48
Query Match          72.9%; Score 146.6; DB 1; Length 542;
Best Local Similarity 91.4%; Pred. No. 8,1e-38;
Matches 149; Conservative 0; Mismatches 14; Indels 0; Gaps 0;
QY 39 AACACTTCTTAACATTACAGAACCTATCCCTGAGAGAGGAGGAGAAAGAACTATTTC 98
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1 AACAAATTTTAAACATTACAGAACGATATCCNTGAGNAGGAGGAGGAGAAAGAAATTTTTC 60
QY 99 ACCCTTGACATGATATTAGTCAAGTCCCTCCCTCTAATTCCTCCATCCCTAGATACAT 158
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 61 ACCCTTGACATGATATTAGTCAAGTCCCTCCCTATATATTCCTCCATCCCTAGANACAT 120
QY 159 CCGGGAAGAGACCTTACCCAGTATTTTATCTACCCCACTGC 201
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 121 CNTGGAGGAGACCTTACCCAGTATTTTATNTATNATACCCCACTGC 163
;
RESULT 3
; Sequence 2, Application US/08721489
; Patent No. 5786465
; GENERAL INFORMATION:
; APPLICANT: Jacobs, Kenneth
; APPLICANT: McCoy, John
; APPLICANT: Lavallee, Edward
; APPLICANT: Racie, Lisa
; APPLICANT: Merberg, David
; APPLICANT: Treacy, Maurice
; APPLICANT: Spaulding, Vikki
```

```

; TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genetics Institute, Inc.
; STREET: 87 Cambridgepark Drive
; CITY: Cambridge
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02140
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/721,489
; FILING DATE:
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Brown, Scott A.
; REGISTRATION NUMBER: 32,724
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 498-8224
; TELEFAX: (617) 876-5851
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 542 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
;
US-08-721-489-2
Query Match          72.9%; Score 146.6; DB 1; Length 542;
Best Local Similarity 91.4%; Pred. No. 8,1e-38;
Matches 149; Conservative 0; Mismatches 14; Indels 0; Gaps 0;
QY 39 AACACTTCTTAACATTACAGAACCTATCCCTGAGAGAGGAGGAGGAGAAAGAACTATTTC 98
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1 AACAAATTTTAAACATTACAGAACGATATCCNTGAGNAGGAGGAGGAGGAGAAAGAAATTTTTC 60
QY 99 ACCCTTGACATGATATTAGTCAAGTCCCTCCCTCTAATTCCTCCATCCCTAGATACAT 158
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 61 ACCCTTGACATGATATTAGTCAAGTCCCTCCCTATATATTCCTCCATCCCTAGANACAT 120
QY 159 CCGGGAAGAGACCTTACCCAGTATTTTATCTACCCCACTGC 201
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 121 CNTGGAGGAGACCTTACCCAGTATTTTATNTATNATACCCCACTGC 163
;
RESULT 4
; Sequence 4, Application US/09146053A
; Patent No. 639349
; GENERAL INFORMATION:
; APPLICANT: Ryan, James W.
; APPLICANT: Spinkie, Jerry Joe Curtis
; APPLICANT: Venema, Richard C.
; TITLE OF INVENTION: Human Aminopeptidase P Gene
; FILE REFERENCE: MCG103
; CURRENT APPLICATION NUMBER: US/09/146,053A
; CURRENT FILING DATE: 1998-09-02
; EARLIER APPLICATION NUMBER: 60/057,854
; EARLIER FILING DATE: 1997-09-02
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 4
; LENGTH: 50000
; TYPE: DNA
; ORGANISM: Homo sapiens
;
US-09-146-053-4
```

Query Match	14.5%;	Score 29.2;	DB 4;	Length 50000;
Best Local Similarity	56.1%;	Pred. No. 8.5;		
Matches 55; Conservative	0;	Mismatches 43;	Indels 0;	Gaps 0;

Df 17852 TGTGATTCCTGTCTTGATGCCAATACCTCTTAATCTGCATCACAGACTAGTCCCTAA 17911

QY 164 GAAGGACCCCTACCCAGATCTTTATCTACCCCACTGC 201
| | | | | | | | | | | | | | | | | | | |
Db 17912 GCCTGGCCACAGACCGAGATCTGTCTTACACACTTGC 17949

RESULT 5
US-09-071-035-63

APPLICANT: Gil H. Choi
 TITLE OF INVENTION: Enterococcus faecalis polynucleotides and polypeptides
 NUMBER OF SEQUENCES: 496

ADDRESSEE: Human Genome Sciences, Inc
STREET: 9410 Key West Avenue
CITY: Rockville
STATE: Maryland
COUNTRY: USA

```

; COMPUTER READABLE FORM:
;
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4mb storage
;
; COMPUTER: HP Vectra 486/33
;
; OPERATING SYSTEM: MSDOS version 6.2
;

```

```

; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/071,035
; FILING DATE:
;

```

;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER:
;; FILING DATE:
;; ATTORNEY/AGENT INFORMATION:

REGISTRATION NUMBER: 36,373
REFERENCE/DOCKET NUMBER: PB369P2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (301) 309-8504

```

; INFORMATION FOR SEQ ID NO: 63
;
; SEQUENCE CHARACTERISTICS:
;     LENGTH: 1564 base pairs
;     TYPE: nucleic acid
;

```

```

; STRANDEDNESS: do
; TOPOLOGY: linear
US-09-071-035-63

```

Query Match	14.2%	Score 28.6;	DB 4;	Length 1564;
Best Local Similarity	59.0%	Pred. No. 4.6;		
Matches 49; Conservative	0;	Mismatches 34;	Indels 0	

Dy 7 AGTGGCAGACATACCCCTCTTAGGCCAAATATCAACAGAAGTTCTTAACAATTGCAAGGAC 66
|| ||| || | | | ||||| || | | ||||| ||| |
Db 1363 ACAAAGCAACGTAACCTTATGCAAATAAATATGAAGAACGTGGAAAAACATTACAAAAGC 1422

Qy 67 CTATCCCTGAGAAGGGAAG 89
||| | | ||| |||
Db 1423 GGATCAATTGTTGCCGAAGAAG 1445

RESULT 6
US-09-071-035-61
; Sequence 61, Application US/09071035
; Patent No. 6448043
; GENERAL INFORMATION:

1 APPLICANT: Gil H. Choi
 2
 3 TITLE OF INVENTION: Enterococcus faecalis polynucleotides and polypeptides
 4
 5 NUMBER OF SEQUENCES: 496
 6
 7 CORRESPONDENCE ADDRESS:
 8
 9 ADDRESSEE: Human Genome Sciences, Inc.

```

; STATE: Maryland
; COUNTRY: USA
;

```

```

;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
;

```

APPLICATION NUMBER: US/09/071, 035
FILING DATE:

;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER:
;; FILING DATE:
;; ATTORNEY/AGENT INFORMATION:

REGISTRATION NUMBER: 36, 373
REFERENCE/DOCKET NUMBER: PB369P2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (301) 309-8504

```

; INFORMATION FOR SEQ ID NO: 61
;
; SEQUENCE CHARACTERISTICS:
;     LENGTH: 1641 base pairs
;     TYPE: nucleic acid

```

TOPLOGY:
US-09-071-035-61

Query Match	14.2%	Score 28.6	DB 4	Length 1641
Best Local Similarly	59.0%	Pred. No. 4.7		
Matches 49; Conservative	0	Mismatches 34	Indels 0	Gaps 0

0y 7 AGTTGGACACATCACTCCCTTAGCCAATATCACAAGTCTTTAAACATTACAGGAC 66
|| ||| | | | ||||| | | | ||||| ||| |
Db 1437 AGAAGCAACGTACTTATTCGAATAATATGAGAAGCGTGAAAAACATTACAAAAAGC 1496

Qy 67 CTATCCCTGAGAAGGGAAAG 89
||| | | | | | |
Db 1497 GGATCAATTGGTTGCCGAAGAAG 1519

RESULT 7
US-08-798-000-3

Patent No. 5766945
GENERAL INFORMATION:
APPLICANT: Miller, A. Dusly
TITLE OF INVENTION: 10A1 RETROVIRAL PACKAGING CELL LINES

;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: Townsend and Crew LLP
;; STREET: TWO Embarcadero Center, 8th Floor
;; CITY: San Francisco

COUNTRY: USA
ZIP: 94111-3834
COMPUTER READABLE

```

: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.25
:
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/798,000
: FILING DATE: 12-FEB-1997
:

```

CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/011,564
FILING DATE: 13-FEB-1996
ATTORNEY/AGENT INFORMATION:
NAME: Parmelee, Steven W.
REGISTRATION NUMBER: 31,990
REFERENCE/DOCKET NUMBER: 14538A-002010
TELECOMMUNICATION INFORMATION:
TELEPHONE: 206-467-9600
TELEFAX: 415-576-0300
INFORMATION FOR SEQ. ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 735 base pairs
TYPE: nucleic acid
STRANDEDNESS: both
TOPOLOGY: linear
MOLECULE TYPE: cDNA
US-08-798-000-3

Query Match 14.1% Score 28.4; DB 1; Length 735;
Best Local Similarity 70.4%; Pred. No. 4.3;
Matches 38; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

QY 148 CCTAGATACATCTGGGAGAGACCTTACCCAGTATTATCTACCCCAACTGC 201
DB 500 CCTAGAACCTCGCTGGAAAGACCTTACACAGTCTGTGACCAACCCCAACGCC 553

RESULT 8

US-08-798-000-1
Sequence 1, Application US/08798000
Patent No. 576945
GENERAL INFORMATION:
APPLICANT: Miller, A. Dusty
TITLE OF INVENTION: 10A1 RETROVIRAL PACKAGING CELL LINES
NUMBER OF SEQUENCES: 3
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, 8th floor
CITY: San Francisco
STATE: CA
COUNTRY: USA
ZIP: 94111-3834
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/798,000
FILING DATE: 12-FEB-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/011,564
FILING DATE: 13-FEB-1996
ATTORNEY/AGENT INFORMATION:
NAME: Parmelee, Steven W.
REGISTRATION NUMBER: 31,990
REFERENCE/DOCKET NUMBER: 14538A-002010
TELECOMMUNICATION INFORMATION:
TELEPHONE: 206-467-9600
TELEFAX: 415-576-0300
INFORMATION FOR SEQ. ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1234 base pairs
TYPE: nucleic acid
STRANDEDNESS: both
TOPOLOGY: linear
MOLECULE TYPE: cDNA
US-08-798-000-1

Query Match 14.1% Score 28.4; DB 1; Length 1234;

Best Local Similarity 70.4%; Pred. No. 5;
Matches 38; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

QY 148 CCTAGATACATCTGGGAGAGACCTTACCCAGTATTATCTACCCCAACTGC 201
DB 500 CCTAGAACCTCGCTGGAAAGACCTTACACAGTCTGTGACCAACCCCAACGCC 553

RESULT 9

US-08-985-526-35
Sequence 35, Application US/08985526
Patent No. 6080728
GENERAL INFORMATION:
APPLICANT: Mixson, James A.
TITLE OF INVENTION: CARRIER-DNA COMPLEXES CONTAINING DNA
TITLE OF INVENTION: ENCODING ANTI-ANGIOGENIC PEPTIDES AND THEIR USE IN GENE
THERAPY
NUMBER OF SEQUENCES: 43
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Connolly, Bove, Lodge, & Hutz
STREET: 1220 Market Street, P.O. Box 2207
CITY: Wilmington
STATE: Delaware
COUNTRY: U.S.A.
ZIP: 19899
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/985,526
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/608,845
FILING DATE: 16-JUL-1996
ATTORNEY/AGENT INFORMATION:
NAME: McMorris Jr., Robert G.
TELECOMMUNICATION INFORMATION:
TELEPHONE: (302) 658-9141
TELEFAX: (302) 658-5613
INFORMATION FOR SEQ. ID NO: 35:
SEQUENCE CHARACTERISTICS:
LENGTH: 2431 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-985-526-35

Query Match 14.1% Score 28.4; DB 3; Length 2431;
Best Local Similarity 54.9%; Pred. No. 6.1;
Matches 56; Conservative 0; Mismatches 46; Indels 0; Gaps 0;

QY 9 TTGCAGACATCATCTCTAGCCCAATATCAACAACTCTTAAACATTACAGAAGACT 68
DB 791 TTGCACACTCTCCACTCAAGTCTCATCAATAGAAAGATTGTAAACCGGATGGAACCC 850
QY 69 ATCCCTGAGAAGAGGAAAGAACTATTCACACCTGTGACA 110
DB 851 TTTCCTGGACACTGTGGCGAAGATGTTTGGACACTTGACA 892

RESULT 10

US-09-265-013-3
Sequence 3, Application US/09265013
Patent No. 6451304
GENERAL INFORMATION:
APPLICANT: FRIEDMAN, Theodore
APPLICANT: MIYANOHARA, Atsushi
TITLE OF INVENTION: METHOD FOR RETROVIRUS VECTOR PRODUCTION BY SEPARATED
FILE REFERENCE: 041673/2010

;; CURRENT APPLICATION NUMBER: US/09/265.013
;; CURRENT FILING DATE: 1999-03-09
;; NUMBER OF SEQ ID NOS: 5
;; SOFTWARE: Patent In Ver. 2.1
;; SEQ ID NO 3
;; LENGTH: 3612
;; TYPE: DNA
;; ORGANISM: Moloney murine leukemia virus (MOLMV)
US-09-265-013-3

Query Match 14.1%; Score 28.4; DB 4; Length 3612;
Best Local Similarity 70.4%; Pred. No. 6.9;
Matches 38; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

OY 148 CCTGATCATCTCGGAGAGACCTACCGCTCTTTACTTACCCACACTGC 201
DB 3414 CCTGACCTCGCTGGAGAGACCTTACACAGCTCTGCTGACACCCACCGC 3467

RESULT 11
US-07-813-593-3
; Sequence 3, Application US/07813593
; Patent No. 5183438
; GENERAL INFORMATION:
; APPLICANT: Lemischka, Ihor R.
; TITLE OF INVENTION: TOTIPOTENT HEMATOPOIETIC STEM CELL
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: IMCONE SYSTEMS INCORPORATED
; STREET: 180 VARICK STREET
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: U.S.A.
; ZIP: 10014
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/813.593
; FILING DATE: 19920415
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/793.065
; FILING DATE: 15-NOV-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/728.913
; FILING DATE: 28-JUN-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/679.666
; FILING DATE: 02-APR-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Feil, Irving N.
; REGISTRATION NUMBER: 28.601
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-645-1405
; TELEFAX: 212-645-2054
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5406 base pairs
; TYPE: NUCLEIC ACID
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 208..4311
; FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: 208..4308

US-07-813-593-3

Query Match 14.1%; Score 28.4; DB 1; Length 5406;
Best Local Similarity 54.9%; Pred. No. 7.8;
Matches 56; Conservative 0; Mismatches 46; Indels 0; Gaps 0;

OY 9 TTGCAGACATCCACTCTCTAGCCAAATATACAGATTTCTTAACATTAAGAAGACCT 68
DB 991 TTGCACCTCTCCACCTTCAAGATCTCATATAGAAAGATTGTAAACCGGAGATGTGAACCC 1050

OY 69 ATCCCTGAGAGAGGAGAAAGACTATTCACCCCTTGACACA 110
DB 1051 TTTCCTGGGACTGTGGCGAAGATGTTTGTGACACCTTGACA 1092

RESULT 12
US-07-977-451-5
; Sequence 5, Application US/07977451
; Patent No. 5270458
; GENERAL INFORMATION:
; APPLICANT: Lemischka, Ihor R.
; TITLE OF INVENTION: TOTIPOTENT HEMATOPOIETIC STEM CELL
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Imclone Systems Incorporated
; STREET: 180 Varick Street
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10014
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/977.451
; FILING DATE: 19921119
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US UNASSIGNED
; FILING DATE: 12-NOV-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/906.397
; FILING DATE: 26-JUN-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US PCT/US92/05401
; FILING DATE: 26-JUN-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: TW 81102961
; FILING DATE: 15-APR-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US PCT/US92/02750
; FILING DATE: 02-APR-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/813.593
; FILING DATE: 24-DEC-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/793.065
; FILING DATE: 15-NOV-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/728.913
; FILING DATE: 28-JUN-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/679.666
; FILING DATE: 02-APR-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Feil, Irving N.
; REGISTRATION NUMBER: 28.601
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-645-1405


```

1      FILING DATE: 02-APR-1992
2      PRIOR APPLICATION DATA:
3      APPLICATION NUMBER: US 07/813,593
4      FILING DATE: 24-DEC-1991
5      PRIOR APPLICATION DATA:
6      APPLICATION NUMBER: US 07/793,065
7      FILING DATE: 15-NOV-1991
8      PRIOR APPLICATION DATA:
9      APPLICATION NUMBER: US 07/728,913
10     FILING DATE: 28-JUN-1991
11     PRIOR APPLICATION DATA:
12     APPLICATION NUMBER: US 07/679,666
13     FILING DATE: 02-APR-1991
14     ATTORNEY/AGENT INFORMATION:
15     NAME: Felt, Irving N.
16     REGISTRATION NUMBER: 28,601
17     REFERENCE/DOCKET NUMBER: LEM-3-7P
18     TELECOMMUNICATION INFORMATION:
19     TELEPHONE: 212-645-1405
20     TELEFAX: 212-645-2054
21     INFORMATION FOR SEQ ID NO: 5:
22     SEQUENCE CHARACTERISTICS:
23     LENGTH: 5406 base pairs
24     TYPE: nucleic acid
25     STRANDEDNESS: double
26     TOPOLOGY: linear
27     MOLECULE TYPE: cdna
28     HYPOTHETICAL: NO
29     ANTI-SENSE: NO
30     FRAGMENT TYPE: N-terminal
31     FEATURE:
32     NAME/KEY: CDS
33     LOCATION: 208..4311
34     FEATURE:
35     NAME/KEY: mat_peptide
36     LOCATION: 265..4308
37     FEATURE:
38     NAME/KEY: sig_peptide
39     LOCATION: 208..264
40     US-08-252-517-5

```

RESULT 15
 US-07-906-397A-5
 : Sequence 5, Application US/07906397A
 : Patent No. 5621090
 : GENERAL INFORMATION:
 : APPLICANT: Lemischka, Ihor R.
 : TITLE OF INVENTION: TOTIPOTENT HEMATOPOIETIC STEM CELL
 : TITLE OF INVENTION: RECEPTORS AND THEIR LIGANDS
 : NUMBER OF SEQUENCES: 6
 : CORRESPONDENCE ADDRESSES:
 : ADDRESSEE: IMCLONE SYSTEMS INCORPORATED
 : STREET: 180 VARICK STREET
 : CITY: NEW YORK
 : STATE: NEW YORK
 : COUNTRY: U.S.A.
 : ZIP: 10014
 : COMPUTER READABLE FORM:
 : MEDIUM TYPE: Floppy disk
 : COMPUTER: IBM PC compatible

```

OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/906,397A
FILING DATE: 19920626
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/813,593
FILING DATE: 24-DEC-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/793,065
FILING DATE: 15-NOV-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/728,913
FILING DATE: 28-JUN-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/679,666
FILING DATE: 02-APR-1991
ATTORNEY/AGENT INFORMATION:
NAME: Felt, Irving N.
REGISTRATION NUMBER: 28,601
REFERENCE/DOCKET NUMBER: LEM-3-PPPPPP
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-645-1405
TELEFAX: 212-645-2054
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 5406 base pairs
TYPE: NUCLEIC ACID
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE: N-terminal
FEATURE:
NAME/KEY: CDS
LOCATION: 208..4311
FEATURE:
NAME/KEY: mat_peptide
LOCATION: 208..4308

```

	Query Match	Score	DB 1	Length	5406
Best Local Similarity	54.9%	Pred.	No. 7.8		
Matches	56	Conservative	0	Mismatches	46
				Indels	0
				Gaps	0
QY	9	TTGCAGACATACCCCTTTAGCCAAATATTCACAAAGTCTCTTAAAAATATTAACGACCT	68		
Db	991	TGCGACCTCTCCACCTTTCAAGTCATCATCAATGAAGATTTGTAACCGGATGTGAAACC	1050		
QY	69	ATTCCTGAGAAAGGAAAAAGACATATTCACCCCTTGACAA	110		
Db	1051	TTTCTTGGAAGACTGTGCGCAAGATGTTTTTGACACACCTTGACA	1092		

Search completed: May 2, 2003, 16:23:51
Job time : 55.3333 secs

GenCore version 5.1.5
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 2, 2003, 16:21:41 ; Search time 63.333 Seconds

(without alignments)
3750.092 Million cell updates/sec

Title: US-09-719-554-3_COPY_7500_7700

Perfect score: 201
Sequence: 1 caactagttgacacatca.....atttaccacccactgc 201

Scoring table:
IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 746064 seqs, 590810554 residues

Total number of hits satisfying chosen parameters: 1492128

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

Published_Applications_NA:*

- 1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
- 2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
- 3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
- 4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
- 5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
- 6: /cgn2_6/ptodata/1/pubpna/PCRTUS_PUBCOMB.seq:*
- 7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
- 8: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
- 9: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
- 10: /cgn2_6/ptodata/1/pubpna/US05_PUBCOMB.seq:*
- 11: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
- 12: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*
- 13: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
- 14: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	201	100.0	2930	10	US-09-902-535-1 Sequence 1, Appl
2	201	100.0	2946	9	US-10-114-893-134 Sequence 134, Appl
3	146.6	72.9	542	12	US-10-040-916-48 Sequence 48, Appl
4	100.8	50.1	758	7	US-08-979-847-112 Sequence 112, Appl
5	46.4	23.1	200	10	US-09-864-761-20979 Sequence 20979, A
6	33.4	16.6	2482	10	US-09-764-853-198 Sequence 198, Appl
7	32.8	16.3	9990	10	US-09-767-088A-2 Sequence 2, Appl
8	32.8	16.3	9990	10	US-09-767-088A-15 Sequence 15, Appl
9	32.6	16.2	10098	9	US-09-764-868-1482 Sequence 1482, Ap
10	31.4	15.6	550	9	US-09-991-836-1452 Sequence 1452, Ap
11	30.8	15.3	509	9	US-09-918-895-11931 Sequence 11931, A
12	30.8	15.3	2456	9	US-09-822-846-42 Sequence 42, Appl
13	30.8	15.3	2496	9	US-10-114-893-122 Sequence 122, Appl
14	30.8	15.3	26197	9	US-10-092-154-1965 Sequence 1965, Ap
15	30.8	15.3	26197	9	US-09-764-847-1965 Sequence 1965, Ap
16	30.8	15.3	26210	9	US-10-092-154-1966 Sequence 1966, Ap
17	30.8	15.3	26210	10	US-09-764-847-1966 Sequence 1966, Ap
18	30.2	15.0	1691139	9	US-10-067-514-1 Sequence 1, Appl
19	30	14.9	1176	10	US-09-764-898-23 Sequence 23, Appl

20	29.6	14.7	159	9	US-09-728-444-16 Sequence 16, Appl
21	29.6	14.7	159	9	US-09-728-444-17 Sequence 17, Appl
22	29.6	14.7	159	9	US-09-728-444-18 Sequence 18, Appl
23	29.6	14.7	159	9	US-09-728-444-96 Sequence 96, Appl
24	29.6	14.7	159	9	US-09-728-444-98 Sequence 98, Appl
25	29.6	14.7	159	9	US-09-728-444-100 Sequence 100, Appl
26	29.6	14.7	258	9	US-09-535-459-1182 Sequence 1182, Appl
27	29.6	14.7	1913	9	US-10-102-806-1 Sequence 1, Appl
28	29.6	14.7	32192	9	US-09-764-891-7049 Sequence 7049, Ap
29	29.4	14.6	432	10	US-09-764-877-2511 Sequence 2511, Ap
30	29.2	14.5	66109	10	US-09-880-107-3768 Sequence 3768, Ap
31	29.2	14.5	106323	10	US-09-880-107-3768 Sequence 3, Appl
32	29.2	14.5	198285	10	US-09-880-107-3814 Sequence 3814, Ap
33	29	14.4	339	9	US-09-803-719-2220 Sequence 2220, Ap
34	29	14.4	446	9	US-10-046-935-1774 Sequence 1774, Ap
35	29	14.4	446	9	US-09-878-178-1774 Sequence 1774, Ap
36	29	14.4	446	9	US-10-146-502-1774 Sequence 1774, Ap
37	29	14.4	501	10	US-09-920-300A-34 Sequence 34, Appl
38	29	14.4	501	12	US-10-033-528-34 Sequence 34, Appl
39	29	14.4	2292	10	US-09-998-598-2602 Sequence 2602, Ap
40	29	14.4	3472	10	US-09-873-737A-5 Sequence 5, Appl
41	28.8	14.3	208	9	US-09-796-692-7291 Sequence 7291, Ap
42	28.8	14.3	208	9	US-10-040-862-7291 Sequence 7291, Ap
43	28.8	14.3	493	10	US-09-783-590-8089 Sequence 8089, Ap
44	28.8	14.3	618	10	US-09-764-898-96 Sequence 96, Appl
45	28.8	14.3	775	9	US-09-796-692-7274 Sequence 7274, Ap

ALIGNMENTS

RESULT 1
US-09-902-535-1
Sequence 1, Application US/09902535
Patent No. US20020102530A1
GENERAL INFORMATION:
APPLICANT: Kelch, Jr., James C.
APPLICANT: McCoy, John M.
TITLE OF INVENTION: Methods and compositions for diagnosing
TITLE OF INVENTION: and treating preclampsia and gestational trophoblast
FILE REFERENCE: GIN-600684
CURRENT APPLICATION NUMBER: US/09/902,535
CURRENT FILING DATE: 2001-07-09
PRIOR APPLICATION NUMBER: 60/216,657
PRIOR FILING DATE: 2000-07-06
NUMBER OF SEQ ID NOS: 5
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 1
LENGTH: 2930
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (930)...(2546)
US-09-902-535-1
Query Match 100.0%; Score 201; DB 10; Length 2930;
Best Local Similarity 100.0%; Pident. No. 5,3e-56;
Matches 201; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CAACTAGTTCGACACATCACCCTCTTGCCCAATATATACAAAGTCTTAACATTACA 60
DB 551 CAACCTAGTTCGACACATCACCCTCTTGCCCAATATATACAAAGTCTTAACATTACA 610
QY 61 AGGAACCATCCTCGAGAGAGGAAAGACATATCCACCTCTGACATGATTTAGT 120
DB 611 AGGAACCATCCTCGAGAGAGGAAAGACATATCCACCTCTGACATGATTTAGT 670
QY 121 CAAGTCCTTCCTCTTAATTCCTCCATCCCTAGATACATCTGGGAAGACCTACCCAGT 180
DB 671 CAAGTCCTTCCTCTTAATTCCTCCATCCCTAGATACATCTGGGAAGACCTACCCAGT 730


```

CORRESPONDENCE ADDRESS:
ADDRESSEE: OLIFF & BERRIDGE, PLC
STREET: P. O. BOX 19928
CITY: ALEXANDRIA
STATE: VA
COUNTRY: USA
ZIP: 22320
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/979,847
FILING DATE: 26-NOV-1997
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: BERRIDGE, WILLIAM P.
REGISTRATION NUMBER: 30,024
REFERENCE/DOCKET NUMBER: MPB 39046A
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-836-6400
TELEFAX: 703-836-2787
INFORMATION FOR SEQ ID NO: 112:
SEQUENCE CHARACTERISTICS:
LENGTH: 758 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
US-08-979-847-112

Query Match          50.1%; Score 100.8; DB 7; Length 758;
Best Local Similarity 93.8%; Pred. No. 2.8e-23;
Matches 105; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 CAACATGATTCGACACATCACCCTCTTACGCAATATCAACAGTCTTAACATTACA 60
    |||||||
Db 647 CAATCTAGTTGACACATCACCCTCTTACGCAATATCAACAGTCTTAACATTACA 706
    |||||||

QY 61 AGAACCCTATCCCTGAGAGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 112
    |||||||
Db 707 GGGAACCTGTCCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 758
    |||||||

RESULT 5
US-09-864-761-20979/c
; Sequence 20979, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Aeomica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; PRIOR FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263,6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
```

```

PRIOR APPLICATION NUMBER: PCT/US01/00664
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00669
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00665
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00668
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00663
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00662
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00661
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00670
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: US 60/234,687
PRIOR FILING DATE: 2000-09-21
PRIOR APPLICATION NUMBER: US 09/608,408
PRIOR FILING DATE: 2000-06-30
PRIOR APPLICATION NUMBER: US 09/774,203
PRIOR FILING DATE: 2001-01-29
NUMBER OF SEQ ID NOS: 49117
SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
SEQ ID NO 20979
LENGTH: 200
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: MAP TO AL034410.8
OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 6.2
OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 1.5
OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.2
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.5
OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.4
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 2.2
OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.7
OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.7
OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.6
OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 1.9
OTHER INFORMATION: NT HIT: g11419444, EVALUATE 5.00e-78
OTHER INFORMATION: EST HUMAN HIT: N76991.1, EVALUATE 1.00e-63
OTHER INFORMATION: SWISSPROT HIT: P03360, EVALUATE 2.00e-06
US-09-864-761-20979

Query Match          23.1%; Score 46.4; DB 10; Length 200;
Best Local Similarity 97.9%; Pred. No. 1.2e-05;
Matches 47; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 153 ATACATCTGTGGAGAGACCTTACCCAGTCATTTATCTACCCCACTG 200
    |||||||
Db 200 ATACATCTGTGGAGAGACCTTACCCAGTCATTTATCTACCCCACTG 153
    |||||||

RESULT 6
US-09-764-853-198/c
; Sequence 198, Application US/09764853
; Patent No. US20020090672A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: Puz06
; CURRENT APPLICATION NUMBER: US/09/764,853
; PRIOR FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 939
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 198
; LENGTH: 2482
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-764-853-198
```

Query Match 16.6%; Score 33.4; DB 10; Length 2482;
Best Local Similarity 54.5%; Pred. No. 0.7;
Matches 67; Conservative 0; Mismatches 56; Indels 0; Gaps 0;

QY 31 CAAATATCAACAAGTTCTTAATAACATTACAGAACCTATCCCTGAGAGGAGAAAGA 90
DB 595 CATAAATGCAAAACCTTTTAAACAAATATAGGAAACCAATCTGGCAACATATTAAGG 536
QY 91 ACTATTCACCCCTTGACATGATTAAGTCAAGTCCCTTCCTCTAATTTCCCATCCCT 150
DB 535 GGTATTACATCATGATGAAAGTGCATTATTCCTGGAATGCCAGGCAATTAATTAATTCAA 476
QY 151 AGA 153
DB 475 AAA 473

RESULT 7

US-09-767-088A-2
; Sequence 2, Application US/09767088A
; Patent No. US20020010947A1
; GENERAL INFORMATION:
; APPLICANT: Gurney, Mark E.
; APPLICANT: Abraham, Irene
; TITLE OF INVENTION: Transgenic Mouse Model Of Human Neurodegenerative Disease
; FILE REFERENCE: PHRM0303
; CURRENT APPLICATION NUMBER: US/09/767,088A
; CURRENT FILING DATE: 2001-01-22
; PRIOR APPLICATION NUMBER: 60/177,319
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: Patentin version 3.0
; SEQ ID NO 2
; LENGTH: 9990
; TYPE: DNA
; ORGANISM: Mus musculus
US-09-767-088A-2

Query Match 16.3%; Score 32.8; DB 10; Length 9990;
Best Local Similarity 59.8%; Pred. No. 2;
Matches 55; Conservative 0; Mismatches 37; Indels 0; Gaps 0;

QY 88 ABAATATTCACACCTTGACATGATTAAGTCAAGTCCCTTCCTCTAATTTCCCATC 147
DB 7772 AACAGTAAACATCTGGAGGTATTAATAGTCTGTGACCCAGAAACCCCACT 7831
QY 148 CCTAGATACATCCTGGAGAGACCTTACCAG 179
DB 7832 CCTGGCAATTTACTGGAGAGAAACAAG 7863

RESULT 8

US-09-767-088A-15
; Sequence 15, Application US/09767088A
; Patent No. US20020010947A1
; GENERAL INFORMATION:
; APPLICANT: Gurney, Mark E.
; APPLICANT: Abraham, Irene
; TITLE OF INVENTION: Transgenic Mouse Model Of Human Neurodegenerative Disease
; FILE REFERENCE: PHRM0303
; CURRENT APPLICATION NUMBER: US/09/767,088A
; CURRENT FILING DATE: 2001-01-22
; PRIOR APPLICATION NUMBER: 60/177,319
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: Patentin version 3.0
; SEQ ID NO 15
; LENGTH: 9990
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: PrP/tau transgene construct

US-09-767-088A-15

Query Match 16.3%; Score 32.8; DB 10; Length 9990;
Best Local Similarity 59.8%; Pred. No. 2;
Matches 55; Conservative 0; Mismatches 37; Indels 0; Gaps 0;

QY 88 ABAATATTCACACCTTGACATGATTAAGTCAAGTCCCTTCCTCTAATTTCCCATC 147
DB 7772 AACAGTAAACATCTGGAGGTATTAATAGTCTGTGACCCAGAAACCCCACT 7831
QY 148 CCTAGATACATCCTGGAGAGACCTTACCAG 179
DB 7832 CCTGGCAATTTACTGGAGAGAAACAAG 7863

RESULT 9

US-09-764-868-1482
; Sequence 1482, Application US/09764868
; Patent No. US2002016871A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PT232
; CURRENT APPLICATION NUMBER: US/09/764,868
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - refer to PAM or file wrapper
; NUMBER OF SEQ ID NOS: 1510
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 1482
; LENGTH: 10098
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-764-868-1482

Query Match 16.2%; Score 32.6; DB 9; Length 10098;
Best Local Similarity 60.9%; Pred. No. 2.3;
Matches 53; Conservative 0; Mismatches 34; Indels 0; Gaps 0;

QY 69 ATCCGTGAGAAAGGAAAGAACTATTCACCCCTTGACATGATTAAGTCAAGTCC 128
DB 1474 AACCTTGAGATTTGGGAAACATATCTCTTTTGTGACATTGATCTGTCTCAT 1533
QY 129 TTCCCTTAATTCCTCATCCCTAGATA 155
DB 1534 TTCTCTCTGTTTCACCAACCTGAATA 1560

RESULT 10

US-09-991-936-1452
; Sequence 1452, Application US/09991936
; Publication No. US20030073827A1
; GENERAL INFORMATION:
; APPLICANT: Brandt, Kevin S.
; APPLICANT: Gaines, Patrick J.
; APPLICANT: Stinchcomb, Dan T.
; APPLICANT: Wisniewski, Nancy
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES, PROTEINS AND USES THEREOF
; FILE REFERENCE: FC-6-cl
; CURRENT APPLICATION NUMBER: US/09/991,936
; CURRENT FILING DATE: 2001-11-21
; PRIOR APPLICATION NUMBER: US/09/543,668
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: 60/128,704
; PRIOR FILING DATE: 1999-04-09
; NUMBER OF SEQ ID NOS: 1959
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 1452
; LENGTH: 550
; TYPE: DNA
; ORGANISM: Ctenocephalides felis
US-09-991-936-1452

Query Match 15.6%; Score 31.4; DB 9; Length 550;
Best Local Similarity 54.9%; Pred. No. 1.7;
Matches 62; Conservative 0; Mismatches 51; Indels 0; Gaps 0;

QY 28 AGCAATATCAACAAGTTCTTAATAACATTCACGAACTATCCCTGAGAGGGGAAA 87
DB 224 AGCCAAATTTTAAAGGCTCTTAATTTAAGAAATATTAATGAGAAACATTTTAA 283

QY 88 AGAATCTTCCACCCCTGTGACATGATATTAAGTACCCCTCCCTCTAAT 140
DB 284 TGAACATTTTAAACGTTAGCAATATTGGAGCTGTGCAATGCTTTAAT 336

RESULT 11

US-09-918-995-11931/c
; Sequence 11931, Application US/09918995
; Publication No. US2003073623A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
; FILE REFERENCE: 20411-756
; CURRENT APPLICATION NUMBER: US/09/918,995
; PRIOR FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: US/09/235,076
; NUMBER OF SEQ ID NOS: 38054
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 11931
; LENGTH: 509
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(509)
; OTHER INFORMATION: n = A,T,C or G
US-09-918-995-11931

Query Match 15.3%; Score 30.8; DB 9; Length 509;
Best Local Similarity 50.7%; Pred. No. 2.5;
Matches 74; Conservative 0; Mismatches 72; Indels 0; Gaps 0;

QY 12 CAGACATCACCTCCTTACCAATATCAACAAGTTCTTAATAACATTCACGAACTATC 71
DB 460 CAGTTCCTCTCCATTTCTGATTTTGTAGATATAAATTTTAAACAAAGGATG 401
QY 72 CCGTGAAGAGGAAAGAAACATATTCACCCCTGTGACATGATTAAGTCAAGTCCCTTC 131
DB 400 TCCCTCTTAGCTAGAGAACTATTCCTCACCTCGTGAAGAGAGAAAAAAGATTC 341

QY 132 CCTCTAATCCCATCCCTAGATACA 157
DB 340 CCATTTCTCTCCTACCTCAAAAACA 315

RESULT 12

US-09-822-846-42/c
; Sequence 42, Application US/09822846
; Publication No. US20030027139A1
; GENERAL INFORMATION:
; APPLICANT: Jacobs, Kenneth
; APPLICANT: McCoy, John M.
; APPLICANT: Lavallee, Edward R.
; APPLICANT: Collins-Racie, Lisa A.
; APPLICANT: Evans, Cheryl
; APPLICANT: Merberg, David
; APPLICANT: Treacy, Maurice
; APPLICANT: Agostino, Michael J.
; APPLICANT: Steininger II, Robert J.
; APPLICANT: Spaulding, Vikki
; APPLICANT: Wong, Gordon G.
; APPLICANT: Clark, Hilary

; APPLICANT: Fechtel, Kim
; APPLICANT: Howes, Steven H.
; APPLICANT: Resnick, Richard J.
; APPLICANT: Gulukota, Kamalakara
; APPLICANT: Graham, James R.
; APPLICANT: Genetics Institute, Inc.
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING NOVEL SECRETED PROTEINS
; FILE REFERENCE: GIN 6400
; CURRENT APPLICATION NUMBER: US/09/822,846
; PRIOR FILING DATE: 2001-03-29
; PRIOR APPLICATION NUMBER: 60/195,605
; PRIOR FILING DATE: 2000-04-06
; NUMBER OF SEQ ID NOS: 629
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 42
; LENGTH: 2456
; TYPE: DNA
; ORGANISM: Homo sapiens

US-09-822-846-42

Query Match 15.3%; Score 30.8; DB 9; Length 2456;
Best Local Similarity 50.7%; Pred. No. 5;
Matches 74; Conservative 0; Mismatches 72; Indels 0; Gaps 0;

QY 13 AGACATCACCTCCTTACCAATATCAACAAGTTCTTAATAACATTCACGAACTATCC 72
DB 634 AATCTCACTTTTCCCAACACACTAAATTTATAGAAACCTAGTAGTACTCTCC 575
QY 73 CTGAGAGAGGAGAAAGAACTATTCACCCCTGTGACATGATTAAGTCAAGTCCCTTC 132
DB 574 ATGACAATGACGTAGAGAAACATCTCAAAACGCTCTCTAGTGTGTCAGGGCTCAAG 515

QY 133 CTCCTAATCCCATCCCTAGATACAT 158
DB 514 ATTGCTTACCAATTTGTCTAAT 489

RESULT 13

US-10-114-893-122/c
; Sequence 122, Application US/10114893
; Publication No. US20020193567A1
; GENERAL INFORMATION:
; APPLICANT: McCoy, John M.
; APPLICANT: Lavallee, Edward R.
; APPLICANT: Collins-Racie, Lisa A.
; APPLICANT: Evans, Cheryl
; APPLICANT: Merberg, David
; APPLICANT: Treacy, Maurice
; APPLICANT: Bowman, Michael R.
; APPLICANT: Spaulding, Vikki
; APPLICANT: Carlin-Duckett, McKeough
; APPLICANT: Kelleher, Kerry S.
; APPLICANT: Genetics Institute, Inc.
; TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES ENCODING THEM
; FILE REFERENCE: GI 6000-10A
; CURRENT APPLICATION NUMBER: US/10/114,893
; PRIOR FILING DATE: 2002-04-02
; EARLIER APPLICATION NUMBER: 09/413,232
; NUMBER OF SEQ ID NOS: 321
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 122
; LENGTH: 2496
; TYPE: DNA
; ORGANISM: Homo sapiens

US-10-114-893-122

Query Match 15.3%; Score 30.8; DB 9; Length 2496;
Best Local Similarity 50.7%; Pred. No. 5;
Matches 74; Conservative 0; Mismatches 72; Indels 0; Gaps 0;

QY 13 AGACATCACCTCCTTACCAATATCAACAAGTTCTTAATAACATTCACGAACTATCC 72

DB 634 AATCTCACTTTTCCCAACACACTAAATTTATAGAAACACTAGTAGATAGCTCTCC 575
QY 73 CTGAGAGAGGGGAAAGAACTATTCACCCCTGTGACATGTTAGTCAAGCCCTTCC 132
DB 574 ATGACAAATGACATGGAAGAACATCTCAAAAGCTCTCTAGTGTGTGTCAGGGCTCAAG 515
QY 133 CTCTAATTCCTCCATCCCTAGATACAT 158
DB 514 ATTTGCTTCACCAAAATTTGTCTAAAT 489

RESULT 14
US-10-092-154-1965
; Sequence 1965, Application US/10092154
; Publication No. US20030054375A1
; GENERAL INFORMATION:

; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC009C1
; CURRENT APPLICATION NUMBER: US/10/092,154
; CURRENT FILING DATE: 2002-03-07
; NUMBER OF SEQ ID NOS: 2003
; Prior Application removed - See File Wrapper or Palm
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO: 1965
; LENGTH: 26197
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-092-154-1965

Query Match 15.3%; Score 30.8; DB 9; Length 26197;
Best Local Similarity 50.7%; Pred. No. 14;
Matches 74; Conservative 0; Mismatches 72; Indels 0; Gaps 0;

QY 13 AGACATCACCCTCTGACCAATATACACAGTCTTAACACTTCAAGAACTATCC 72
DB 3270 AATCTCACTTTTCCCAACACACTAAATTTATAGAAACACTAGTAGATAGCTCTCC 3329
QY 73 CTGAGAGAGGGGAAAGAACTATTCACCCCTGTGACATGTTAGTCAAGTCCCTTCC 132
DB 3330 ATGACAAATGACATGGAAGAACATCTCAAAAGCTCTCTAGTGTGTGTCAGGGCTCAAG 3389
QY 133 CTCTAATTCCTCCATCCCTAGATACAT 158
DB 3390 ATTTGCTTCACCAAAATTTGTCTAAAT 3415

RESULT 15
US-09-764-847-1965
; Sequence 1965, Application US/09764847
; Patent No. US20020132767A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC009
; CURRENT APPLICATION NUMBER: US/09/764,847
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 2003
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO: 1965
; LENGTH: 26197
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-764-847-1965

Query Match 15.3%; Score 30.8; DB 10; Length 26197;
Best Local Similarity 50.7%; Pred. No. 14;
Matches 74; Conservative 0; Mismatches 72; Indels 0; Gaps 0;

QY 13 AGACATCACCCTCTGACCAATATACACAGTCTTAACACTTCAAGAACTATCC 72
DB 3270 AATCTCACTTTTCCCAACACACTAAATTTATAGAAACACTAGTAGATAGCTCTCC 3329
QY 73 CTGAGAGAGGGGAAAGAACTATTCACCCCTGTGACATGTTAGTCAAGTCCCTTCC 132
DB 3330 ATGACAAATGACATGGAAGAACATCTCAAAAGCTCTCTAGTGTGTGTCAGGGCTCAAG 3389
QY 133 CTCTAATTCCTCCATCCCTAGATACAT 158
DB 3390 ATTTGCTTCACCAAAATTTGTCTAAAT 3415

DB 3270 AATCTCACTTTTCCCAACACACTAAATTTATAGAAACACTAGTAGATAGCTCTCC 3329
QY 73 CTGAGAGAGGGGAAAGAACTATTCACCCCTGTGACATGTTAGTCAAGTCCCTTCC 132
DB 3330 ATGACAAATGACATGGAAGAACATCTCAAAAGCTCTCTAGTGTGTGTCAGGGCTCAAG 3389
QY 133 CTCTAATTCCTCCATCCCTAGATACAT 158
DB 3390 ATTTGCTTCACCAAAATTTGTCTAAAT 3415

Search completed: May 2, 2003, 18:18:23
Job time : 86.3333 secs

GenCore version 5.1.5
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 2, 2003, 12:54:45 ; Search time 689 Seconds

(without alignments)
8490.078 Million cell updates/sec

Title: US-09-719-554-3_COPY_8500_8700

Perfect score: 201
Sequence: 1 aactcagcagcaagaataaa.....tatttttgcgtgtracc 201

Scoring table: IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl: *
1: gb_da: *
2: gb_htg: *
3: gb_in: *
4: gb_ov: *
5: gb_ov: *
6: gb_pat: *
7: gb_ph: *
8: gb_pl: *
9: gb_pr: *
10: gb_ro: *
11: gb_sts: *
12: gb_sy: *
13: gb_un: *
14: gb_vi: *
15: em_ba: *
16: em_fun: *
17: em_hum: *
18: em_in: *
19: em_mu: *
20: em_or: *
21: em_om: *
22: em_ov: *
23: em_pat: *
24: em_ph: *
25: em_pl: *
26: em_ro: *
27: em_sts: *
28: em_un: *
29: em_vi: *
30: em_htg_hum: *
31: em_htg_inv: *
32: em_htg_other: *
33: em_htg_mus: *
34: em_htg_pln: *
35: em_htg_rnd: *
36: em_htg_mam: *
37: em_htg_vrl: *
38: em_sy: *
39: em_htgo_hum: *
40: em_htgo_mus: *
41: em_htgo_other: *

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	201	100.0	540	9 AB051004	AB051004 Homo sapi
2	201	100.0	540	9 AB079501	AB079501 Homo sapi
3	201	100.0	541	9 AB079503	AB079503 Homo sapi
4	201	100.0	541	9 AB079504	AB079504 Homo sapi
5	201	100.0	541	9 AB079509	AB079509 Homo sapi
6	201	100.0	542	9 AB079502	AB079502 Homo sapi
7	201	100.0	543	9 AB054089	AB054089 Homo sapi
8	201	100.0	543	9 AB055416	AB055416 Homo sapi
9	201	100.0	1617	9 AF156963	AF156963 Homo sapi
10	201	100.0	1624	9 AF506835	AF506835 Homo sapi
11	201	100.0	1860	9 AF513360	AF513360 Homo sapi
12	201	100.0	1948	6 AX000960	AX000960 Sequence
13	201	100.0	1948	6 AX027474	AX027474 Sequence
14	201	100.0	1948	9 AF072505	AF072505 Homo sapi
15	201	100.0	2055	6 AX007999	AX007999 Sequence
16	201	100.0	2599	6 AX007978	AX007978 Sequence
17	201	100.0	2781	6 AX092223	AX092223 Sequence
18	201	100.0	2781	6 AF072506	AF072506 Homo sapi
19	201	100.0	2930	6 AX355872	AX355872 Sequence
20	201	100.0	2930	9 AF208161	AF208161 Homo sapi
21	201	100.0	2946	6 AR177269	AR177269 Sequence
22	201	100.0	10499	6 AX007980	AX007980 Sequence
23	201	100.0	56093	6 AX329572	AX329572 Sequence
24	201	100.0	56093	9 HSAC000064	HSAC000064 Human BAC
25	201	100.0	149194	9 AC007566	AC007566 Homo sapi
26	200.6	99.8	7582	6 AX000966	AX000966 Sequence
27	200.6	99.8	7582	6 AX027480	AX027480 Sequence
28	199.4	99.2	2782	6 AX000962	AX000962 Sequence
29	199.4	99.2	2782	6 AX027476	AX027476 Sequence
30	199.4	99.2	2782	6 AX133396	AX133396 Sequence
31	194.6	96.8	540	9 AB055412	AB055412 Homo sapi
32	193	96.0	540	9 AB051010	AB051010 Homo sapi
33	193	96.0	1481	6 AX001027	AX001027 Sequence
34	193	96.0	1629	6 AX036992	AX036992 Sequence
35	193	96.0	1629	14 AF331500	AF331500 Multiple
36	193	96.0	1799	6 AX007982	AX007982 Sequence
37	193	96.0	1932	14 AF127228	AF127228 Multiple
38	193	96.0	46340	6 AX008043	AX008043 Sequence
39	193	96.0	128485	2 AL358814	AL358814 Homo sapi
40	193	96.0	162579	9 AL390039	AL390039 Human DNA
41	193	96.0	172918	2 AC023366	AC023366 Homo sapi
42	193	95.2	177163	9 HS134E15	AL022067 Human DNA
43	191.4	95.2	539	9 AB054086	AB054086 Homo sapi
44	191.4	95.2	540	9 AB051003	AB051003 Homo sapi
45	191.4	95.2	540	9 AB051007	AB051007 Homo sapi

ALIGNMENTS

RESULT 1	AB051004	540 bp	DNA	linear	PRI 18-MAY-2001
LOCUS	AB051004				
DEFINITION	Homo sapiens human endogenous retrovirus W gene for envelope protein, partial cds, isolate:W-7-1.				
ACCESSION	AB051004				
VERSION	AB051004.1	GI:14020829			
KEYWORDS					
SOURCE					
ORGANISM	Homo sapiens (isolate:W-7-1) DNA.				
REFERENCE					
AUTHORS	Kim,H.S. and Lee,W.H.				
TITLE	Human endogenous retrovirus HERV-W family: chromosomal				

Pred. No. is the number of results predicted by chance to have a

JOURNAL
MEDLINE
REFERENCE
AUTHORS
JOURNAL
TITLE
Direct Submission
Submitted (10-NOV-2000) Heui-Seo Kim, Pusan National University,
Biology; Keumjeong Ku, Changjeon Dong, Pusan 609-735, Korea
(E-mail: khs307@hyowon.cc.pusan.ac.kr, URL: http://mpl.biology.or.kr,
Tel: 82515102259, Fax: 82515812962)

localization, identification, and phylogeny
AIDS Res. Hum. Retroviruses 17 (7), 643-648 (2001)
21268768
2 (bases 1 to 540)
Kim, H.
Location/Qualifiers
1..540
/organism="Homo sapiens"
/isolate="W-7-1"
/db_xref="taxon:9606"
/chromosome="7"
1..540
/gene="env"
<1..>540
/gene="env"
/codon_start=1
/product="envelope protein"
/protein_id="BAB47558.1"
/db_xref="GI:14020830"
/translation="NNEPHEINTSVLVGPIVSNLEIPIHTSNLGCVRKSNITYTTNSO
CIRWTPPTQIVCLPSGIFVCGTSAYRGLNSSSMCLSPVLPMTYTRDQYXY
VTSKPRNRVPLIPVIGAGVIGALGTGIGTITSTQPIYKLSQELNGMERYADSLV
TLQDNLSTAAVVLONRRAL"

gene
CDS

repeat_region
1..540
/rpt_family="HERV-W"
BASE COUNT 156 a 146 c 89 g 149 t
ORIGIN

Query Match
Best Local Similarity 100.0%; Score 201; DB 9; Length 540;
Matches 201; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AACTCAGCAGAGAAATAAACACCACTCCGTTTGTAGTAGGACCTCTGTTTCCAATCTG 60
|||||
DB 4 AACTCAGCAGAGAAATAAACACCACTCCGTTTGTAGTAGGACCTCTGTTTCCAATCTG 63
QY 61 GAATATACCATCTCAACCTCACTGCTGTAAATTAGCAATCTACATACACAACC 120
|||||
DB 64 GAATATACCATCTCAACCTCACTGCTGTAAATTAGCAATCTACATACACAACC 123
QY 121 AACTCCCAATGATCAGGTGGGTACTCTCCACACAATACTGCTACCTCAGGA 180
|||||
DB 124 AACTCCCAATGATCAGGTGGGTACTCTCCACACAATACTGCTACCTCAGGA 183
QY 181 ATATTTTGTCTGTGTGTACC 201
|||||
DB 184 ATATTTTGTCTGTGTGTACC 204

RESULT 2
AB079501
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Homo sapiens (isolate:HB-1) brain CDNA to mRNA.
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
Shin, K., Yi, J. and Kim, H.
Isolation and Phylogeny of HERV-W env Fragments in Human Brain CDNA
library
Unpublished
2 (bases 1 to 540)
Shin, K., Yi, J. and Kim, H.
Direct Submission

JOURNAL
Submitted (05-FEB-2002) Heui-Seo Kim, Pusan National University,
Biology; 30 Changjeon-dong Keumjeong-ku, Pusan 609-735, Republic
of Korea (E-mail: khs307@hyowon.cc.pusan.ac.kr, Tel: 82515103346,
Fax: 82515812962)

FEATURES
source
1..540
/organism="Homo sapiens"
/isolate="HB-1"
/db_xref="taxon:9606"
/tissue_type="brain"
repeat_region
1..540
/rpt_family="HERV-W"
BASE COUNT 156 a 146 c 89 g 149 t
ORIGIN

Query Match
Best Local Similarity 100.0%; Score 201; DB 9; Length 540;
Matches 201; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AACTCAGCAGAGAAATAAACACCACTCCGTTTGTAGTAGGACCTCTGTTTCCAATCTG 60
|||||
DB 4 AACTCAGCAGAGAAATAAACACCACTCCGTTTGTAGTAGGACCTCTGTTTCCAATCTG 63
QY 61 GAATATACCATCTCAACCTCACTGCTGTAAATTAGCAATCTACATACACAACC 120
|||||
DB 64 GAATATACCATCTCAACCTCACTGCTGTAAATTAGCAATCTACATACACAACC 123
QY 121 AACTCCCAATGATCAGGTGGGTACTCTCCACACAATACTGCTACCTCAGGA 180
|||||
DB 124 AACTCCCAATGATCAGGTGGGTACTCTCCACACAATACTGCTACCTCAGGA 183
QY 181 ATATTTTGTCTGTGTGTACC 201
|||||
DB 184 ATATTTTGTCTGTGTGTACC 204

RESULT 3
AB079503
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Homo sapiens (isolate:HB-3) brain CDNA to mRNA.
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
Shin, K., Yi, J. and Kim, H.
Isolation and Phylogeny of HERV-W env Fragments in Human Brain CDNA
library
Unpublished
2 (bases 1 to 541)
Shin, K., Yi, J. and Kim, H.
Direct Submission
Submitted (05-FEB-2002) Heui-Seo Kim, Pusan National University,
Biology; 30 Changjeon-dong Keumjeong-ku, Pusan 609-735, Republic
of Korea (E-mail: khs307@hyowon.cc.pusan.ac.kr, Tel: 82515103346,
Fax: 82515812962)

FEATURES
source
1..541
/organism="Homo sapiens"
/isolate="HB-3"
/db_xref="taxon:9606"
/tissue_type="brain"
repeat_region
1..541
/rpt_family="HERV-W"
BASE COUNT 154 a 146 c 92 g 149 t
ORIGIN

Query Match
Best Local Similarity 100.0%; Score 201; DB 9; Length 541;
Matches 201; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

OY      1  AACCTCAGCAGAAATAAACCACCTCGTTTAGTAGGACCTGTTTCCAAATCG 60
        |||||||
DB      4  AACCTCAGCAGAAATAAACCACCTCGTTTAGTAGGACCTGTTTCCAAATCG 63
        |||||||
OY      61  GAAATTAACCCATACCTCAACCTCAGCTGTGTAAATTTAGCAATACATACACCAAC 120
        |||||||
DB      64  GAAATTAACCCATACCTCAACCTCAGCTGTGTAAATTTAGCAATACATACACCAAC 123
        |||||||
OY      121  AACCTCCCAATGCATGAGTGGGTAACTCTCCACACAAATAGTCTGCCCTACCTCAGGA 180
        |||||||
DB      124  AACCTCCCAATGCATGAGTGGGTAACTCTCCACACAAATAGTCTGCCCTACCTCAGGA 183
        |||||||
OY      181  ATATTTTGTCTGTGTGTAC 201
        |||||||
DB      184  ATATTTTGTCTGTGTGTAC 204
        |||||||

RESULT 4
AB079504      541 bp  mRNA  linear  PRI 06-AUG-2002
LOCUS
DEFINITION  Homo sapiens mRNA, repeat sequence HERV-W, isolate:HB-4.
ACCESSION  AB079504
VERSION  AB079504.1 GI:22122078
KEYWORDS
SOURCE
ORGANISM  Homo sapiens (isolate:HB-4) brain cDNA to mRNA.
           Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
           Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
AUTHORS  1  Shin, K., Yi, J. and Kim, H.
TITLE  Isolation and Phylogeny of HERV-W env Fragments in Human Brain cDNA
        library
JOURNAL  Unpublished
REFERENCE  2  (bases 1 to 541)
AUTHORS  Shin, K., Yi, J. and Kim, H.
TITLE  Direct Submission
JOURNAL  Submitted (05-FEB-2002) Heui-Soo Kim, Pusan National University,
        Biology: 30 Changjeon-dong Keumjeoung-ku, Pusan 609-735, Republic
        of Korea (E-mail:kims307@hyowon.cc.pusan.ac.kr, Tel:82515103346,
        Fax:8251812962)
        Location/Qualifiers
FEATURES
source
1..541
/organism="Homo sapiens"
/isolate="HB-4"
/db_xref="taxon:9606"
/tissue_type="brain"
repeat_region 1..541
BASE COUNT  155 a 144 c 90 g 152 t
ORIGIN
Query Match 100.0%; Score 201; DB 9; Length 541;
Best Local Similarity 100.0%; Pred. No. 1,1e-48;
Matches 201; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY      1  AACCTCAGCAGAAATAAACCACCTCGTTTAGTAGGACCTGTTTCCAAATCG 60
        |||||||
DB      4  AACCTCAGCAGAAATAAACCACCTCGTTTAGTAGGACCTGTTTCCAAATCG 63
        |||||||
OY      61  GAAATTAACCCATACCTCAACCTCAGCTGTGTAAATTTAGCAATACATACACCAAC 120
        |||||||
DB      64  GAAATTAACCCATACCTCAACCTCAGCTGTGTAAATTTAGCAATACATACACCAAC 123
        |||||||
OY      121  AACCTCCCAATGCATGAGTGGGTAACTCTCCACACAAATAGTCTGCCCTACCTCAGGA 180
        |||||||
DB      124  AACCTCCCAATGCATGAGTGGGTAACTCTCCACACAAATAGTCTGCCCTACCTCAGGA 183
        |||||||
OY      181  ATATTTTGTCTGTGTGTAC 201
        |||||||
DB      184  ATATTTTGTCTGTGTGTAC 204
        |||||||
```

```

RESULT 5
AB079509      541 bp  mRNA  linear  PRI 06-AUG-2002
LOCUS
DEFINITION  Homo sapiens mRNA, repeat sequence HERV-W, isolate:HB-9.
ACCESSION  AB079509
VERSION  AB079509.1 GI:22122083
KEYWORDS
SOURCE
ORGANISM  Homo sapiens (isolate:HB-9) brain cDNA to mRNA.
           Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
           Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
AUTHORS  1  Shin, K., Yi, J. and Kim, H.
TITLE  Isolation and Phylogeny of HERV-W env Fragments in Human Brain cDNA
        library
JOURNAL  Unpublished
REFERENCE  2  (bases 1 to 541)
AUTHORS  Shin, K., Yi, J. and Kim, H.
TITLE  Direct Submission
JOURNAL  Submitted (05-FEB-2002) Heui-Soo Kim, Pusan National University,
        Biology: 30 Changjeon-dong Keumjeoung-ku, Pusan 609-735, Republic
        of Korea (E-mail:kims307@hyowon.cc.pusan.ac.kr, Tel:82515103346,
        Fax:8251812962)
        Location/Qualifiers
FEATURES
source
1..541
/organism="Homo sapiens"
/isolate="HB-9"
/db_xref="taxon:9606"
/tissue_type="brain"
repeat_region 1..541
BASE COUNT  156 a 146 c 89 g 150 t
ORIGIN
Query Match 100.0%; Score 201; DB 9; Length 541;
Best Local Similarity 100.0%; Pred. No. 1,1e-48;
Matches 201; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY      1  AACCTCAGCAGAAATAAACCACCTCGTTTAGTAGGACCTGTTTCCAAATCG 60
        |||||||
DB      4  AACCTCAGCAGAAATAAACCACCTCGTTTAGTAGGACCTGTTTCCAAATCG 63
        |||||||
OY      61  GAAATTAACCCATACCTCAACCTCAGCTGTGTAAATTTAGCAATACATACACCAAC 120
        |||||||
DB      64  GAAATTAACCCATACCTCAACCTCAGCTGTGTAAATTTAGCAATACATACACCAAC 123
        |||||||
OY      121  AACCTCCCAATGCATGAGTGGGTAACTCTCCACACAAATAGTCTGCCCTACCTCAGGA 180
        |||||||
DB      124  AACCTCCCAATGCATGAGTGGGTAACTCTCCACACAAATAGTCTGCCCTACCTCAGGA 183
        |||||||
OY      181  ATATTTTGTCTGTGTGTAC 201
        |||||||
DB      184  ATATTTTGTCTGTGTGTAC 204
        |||||||

RESULT 6
AB079502      542 bp  mRNA  linear  PRI 06-AUG-2002
LOCUS
DEFINITION  Homo sapiens mRNA, repeat sequence HERV-W, isolate:HB-2.
ACCESSION  AB079502
VERSION  AB079502.1 GI:22122076
KEYWORDS
SOURCE
ORGANISM  Homo sapiens (isolate:HB-2) brain cDNA to mRNA.
           Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
           Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
AUTHORS  1  Shin, K., Yi, J. and Kim, H.
TITLE  Isolation and Phylogeny of HERV-W env Fragments in Human Brain cDNA
        library
JOURNAL  Unpublished
REFERENCE  2  (bases 1 to 542)
AUTHORS  Shin, K., Yi, J. and Kim, H.
```

```

TITLE      Direct Submission
JOURNAL    Submitted (05-FEB-2002) Heui-Soo Kim, Pusan National University,
            Biology; 30 Changjeon-dong Keunjeoung-ku, Pusan 609-735, Republic
            of Korea (E-mail:khs307ehyowon.cc.pusan.ac.kr, Tel:82515103346,
            Fax:82515812962)
FEATURES   source
            1..542
            /organism="Homo sapiens"
            /isolate="HB-2"
            /db_xref="taxon:9606"
            /tissue_type="brain"
            repeat_region 1..542
            /rpt_family="HERV-W"
BASE COUNT 157 a 145 c 91 g 149 t
ORIGIN
Query Match      100.0%; Score 201; DB 9; Length 542;
Best Local Similarity 100.0%; Pred. No. 1.le-48;
Matches 201; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AACTGAGCAGCAAAATAAACACCACTCCGTTTGTAGTAGACCTCTGTTCCATCTG 60
    |||||||
DB 4 AACTGAGCAGCAAAATAAACACCACTCCGTTTGTAGTAGACCTCTGTTCCATCTG 63
QY 61 GAAATTAACCCATACCTCAAACTCACCCTGTGTAAATTTAGCAATACATACACAAACC 120
    |||||||
DB 64 GAAATTAACCCATACCTCAAACTCACCCTGTGTAAATTTAGCAATACATACACAAACC 123
QY 121 AACTCCCAATGATCAGTGGGTGTTACTCTCCACACAAATAGTGTGCTTACCTCAGGA 180
    |||||||
DB 124 AACTCCCAATGATCAGTGGGTGTTACTCTCCACACAAATAGTGTGCTTACCTCAGGA 183
QY 181 ATATTTTGTCTGTGTGTACC 201
    |||||||
DB 184 ATATTTTGTCTGTGTGTACC 204

RESULT 7
LOCUS      AB054089 543 bp DNA linear PRI 06-DEC-2001
DEFINITION Homo sapiens DNA, human endogenous retrovirus HERV-W family
ACCESSION  AB054089
VERSION     AB054089.1 GI:14780053
KEYWORDS
SOURCE      Homo sapiens (isolate:UO31-3) cell_line:UO-31 DNA.
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
REFERENCE   1
            Yl,J.-M., Kim,H.-M., Lee,W.-H. and Kim,H.-S.
            Molecular cloning and phylogenetic analysis of new human endogenous
            retrovirus HERV-W family in cancer cells
            JOURNAL Curr. Microbiol. (2001) In press
            AUTHORS Lee,J. and Kim,H.-S.
            TITLE 2 (bases 1 to 543)
            JOURNAL Direct Submission
            SUBMITTED (15-JAN-2001) Joo-Mi Lee, Pusan National University,
            Biology; 30 Changjeon-dong Keunjeoung-ku, Pusan 609-735, Republic
            of Korea (E-mail:jmllee76@hotmail.com, Tel:82515103346,
            Fax:82515812962)
FEATURES   Location/Qualifiers
            source 1..543
            /organism="Homo sapiens"
            /isolate="UO31-3"
            /db_xref="taxon:9606"
            /cell_line="UO-31"
            /note="HERV-W env"
            repeat_region 1..543
            /rpt_family="human endogenous retrovirus HERV-W"
BASE COUNT 156 a 146 c 89 g 152 t
ORIGIN
Query Match      100.0%; Score 201; DB 9; Length 543;
Best Local Similarity 100.0%; Pred. No. 1.le-48;
Matches 201; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AACTGAGCAGCAAAATAAACACCACTCCGTTTGTAGTAGACCTCTGTTCCATCTG 60
    |||||||
DB 4 AACTGAGCAGCAAAATAAACACCACTCCGTTTGTAGTAGACCTCTGTTCCATCTG 63
QY 61 GAAATTAACCCATACCTCAAACTCACCCTGTGTAAATTTAGCAATACATACACAAACC 120
    |||||||
DB 64 GAAATTAACCCATACCTCAAACTCACCCTGTGTAAATTTAGCAATACATACACAAACC 123
QY 121 AACTCCCAATGATCAGTGGGTGTTACTCTCCACACAAATAGTGTGCTTACCTCAGGA 180
    |||||||
DB 124 AACTCCCAATGATCAGTGGGTGTTACTCTCCACACAAATAGTGTGCTTACCTCAGGA 183
QY 181 ATATTTTGTCTGTGTGTACC 201
    |||||||
DB 184 ATATTTTGTCTGTGTGTACC 204

RESULT 8
LOCUS      AB055416 543 bp DNA linear PRI 07-DEC-2001
DEFINITION Homo sapiens DNA, human endogenous retrovirus HERV-W family
ACCESSION  AB055416
VERSION     AB055416.1 GI:15128498
KEYWORDS
SOURCE      Homo sapiens (isolate:RT4-2) cell_line:RT4 DNA.
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
REFERENCE   1
            Yl,J.-M., Lee,W.-H., Kim,H.-M. and Kim,H.-S.
            Identification of new endogenous retroviral sequences belonging to
            the HERV-W family in human cancer cells
            JOURNAL Intervirology (2001) In press
            AUTHORS Lee,J. and Kim,H.
            TITLE 2 (bases 1 to 543)
            JOURNAL Direct Submission
            SUBMITTED (04-FEB-2001) Joo-Mi Lee, Pusan National University,
            Biology; 30 Changjeon-dong Keunjeoung-ku, Pusan 609-735, Republic
            of Korea (E-mail:jmllee76@hotmail.com, Tel:82515103346,
            Fax:82515812962)
FEATURES   Location/Qualifiers
            source 1..543
            /organism="Homo sapiens"
            /isolate="RT4-2"
            /db_xref="taxon:9606"
            /cell_line="RT4"
            /note="HERV-W env"
            repeat_region 1..543
            /rpt_family="human endogenous retrovirus HERV-W"
BASE COUNT 156 a 146 c 89 g 152 t
ORIGIN
Query Match      100.0%; Score 201; DB 9; Length 543;
Best Local Similarity 100.0%; Pred. No. 1.le-48;
Matches 201; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AACTGAGCAGCAAAATAAACACCACTCCGTTTGTAGTAGACCTCTGTTCCATCTG 60
    |||||||
DB 4 AACTGAGCAGCAAAATAAACACCACTCCGTTTGTAGTAGACCTCTGTTCCATCTG 63
QY 61 GAAATTAACCCATACCTCAAACTCACCCTGTGTAAATTTAGCAATACATACACAAACC 120
    |||||||
DB 64 GAAATTAACCCATACCTCAAACTCACCCTGTGTAAATTTAGCAATACATACACAAACC 123
QY 121 AACTCCCAATGATCAGTGGGTGTTACTCTCCACACAAATAGTGTGCTTACCTCAGGA 180
    |||||||
DB 124 AACTCCCAATGATCAGTGGGTGTTACTCTCCACACAAATAGTGTGCTTACCTCAGGA 183
QY 181 ATATTTTGTCTGTGTGTACC 201
    |||||||
DB 184 ATATTTTGTCTGTGTGTACC 204
```

QY 181 ATATTTTGTCTGTGTACC 201
|||||
Db 184 ATATTTTGTCTGTGTACC 204

RESULT 9
AF156963 1617 bp DNA linear PRI 06-JUN-2000
LOCUS Homo sapiens human endogenous retrovirus W envC7-1 envelope protein
DEFINITION (env) gene, complete cds.
ACCESSION AF156963
VERSION AF156963.1 GI:8272467
KEYWORDS
SOURCE Homo sapiens.
ORGANISM Homo sapiens.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 1617)
Voisset,C., Bouton,O., Bedin,F., Duret,L., Mandrand,B., Mallet,F.
and Paranhos-Baccala,G.
Chromosomal distribution and coding capacity of the human
endogenous retrovirus HERV-W family
AIDS Res. Hum. Retroviruses 16 (8), 731-740 (2000)
JOURNAL MEDLINE
20284713
PUBMED 10826480
REFERENCE 2 (bases 1 to 1617)
Voisset,C., Bouton,O., Bedin,F., Duret,L., Mandrand,B., Mallet,F.
and Paranhos-Baccala,G.
TITLE Direct Submission
JOURNAL Submitted (09-JUN-1999) UMR103CNRS-Biomerieux, ENS Lyon, 46 allée
d'Italie, Lyon 69364, France
FEATURES
source
1..1617
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="envC7-1"
/note="human/rodent monochromosomal NIGMS somatic hybrid
panel #2 (Coriell Institute, Camden, NJ, USA); clone
GM/NA10791 retaining human chromosome 7"
1..1617
/gene="env"
/gene="env"
1..1617
/product="envelope protein"
/protein_id="AA074215.1"
/db_xref="GI:8272468"
/translation="MALPYHIFLFTVLLPSFTLLAPPCRCMTSSSPYOEFLMORP
GNIDAPYSRSLSKGTPPTATHTHPRNCYHSATLCMAHTHYTGKMINPSCPGGLV
TVCWTYFTQGMDSGGVODARREKHVEVLSQTRVHGTSPPYKGLDLSLHETLRT
HTRLVSLFNTLLGLHEVSAQNPNCWICLPINRPVYSIPVPOMNNFSEINTTSV
LVGLVSNLEITHRSNLTVCVFSNTYTTNSQICRWYTPPOIVCLPSGIFVCGTSA
YRCLNGSSEMCFLSPVPMPTIYTEDOLYVSKPRNKRPVILPVIAGVIGALG
TGIGITTSQFYKLSQELNGDMERVADSVTLQDOLNSIAAVLGNRRALDLITAE
RGGTCFLGEECCYVNOGSLVTEKVEIRRIORRAELRNTPGMGLISQMMWILP
FLGPLAIIILLFLGPCIFNLVNFVSSRIEAVKLQMPKQSTKTKYRRPLDRPAS
RSDVNDIKGTPPEIISAQPLRNAGSS"
repeat_region
1..1617
/rpt_family="human endogenous retrovirus W"
/rpt_type="dispersed"
1..18
/gene="env"
1597..1617
/gene="env"
BASE COUNT 441 a 452 c 297 g 427 t
ORIGIN

Query Match 100.0%; Score 201; DB 9; Length 1617;
Best Local Similarity 100.0%; Pred. No. 1e-48;
Matches 201; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AACTTCAGCAGCAAGAAATTAACACACCTCCGTTTAGTAGAGACCTCTGTTCCATCTG 60

Db 622 AACTTCAGCAGCAAGAAATTAACACACCTCCGTTTAGTAGAGACCTCTGTTCCATCTG 681
|||||
QY 61 GAATTAACCATCTACTCAAACTCAGCTGTGTGTAATTTAGCATACTACATCAAC 120
|||||
Db 682 GAATTAACCATCTACTCAAACTCAGCTGTGTGTAATTTAGCATACTACATCAAC 741
|||||

QY 121 AACTCCCAATGATGAGGAGGAGTACCTCTCCACACAATAGTGTGCTACCCACAGA 180
|||||
Db 742 AACTCCCAATGATGAGGAGGAGTACCTCTCTCCACACAATAGTGTGCTACCCACAGA 801
|||||

QY 181 ATATTTTGTCTGTGTACC 201
|||||
Db 802 ATATTTTGTCTGTGTACC 822

RESULT 10
AF506835 1624 bp mRNA linear PRI 20-MAY-2002
LOCUS Homo sapiens envyrl mRNA, partial cds.
DEFINITION AF506835
ACCESSION AF506835.1 GI:20978307
KEYWORDS
SOURCE Homo sapiens.
ORGANISM Homo sapiens.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 1624)
Alliel,P.M., Perin,J.P., Pierig,R., Nussbaum,J.L., Menard,A. and
Rieger,F.
Endogenous retroviruses and multiple sclerosis. II. HERV-7q
C. R. Acad. Sci. III, Sci. Vie 321 (10), 857-863 (1998)
JOURNAL MEDLINE
96052087
PUBMED 9835022
REFERENCE 2 (bases 1 to 1624)
Alliel,P.M., Perin,J.P., Goudou,D., Bitoun,M., Robert,B. and
Rieger,F.
The HERV-W/7q family in the human genome. Potential for protein
expression and gene regulation
Cell. Mol. Biol. 48, 213-217 (2002)
JOURNAL 3 (bases 1 to 1624)
Goudou,D., Perin,J.P., Rieger,F., Robert,B. and Alliel,P.M.
TITLE Direct Submission
JOURNAL Submitted (29-APR-2002) U-488, INSERM, 80, rue du General Leclerc,
Le Kremlin-Bicetre 94270, France
FEATURES
source
1..1624
/organism="Homo sapiens"
/db_xref="taxon:9606"
/tissue="placenta"
25..>1624
/codon_start=1
/product="envyrl"
/protein_id="AA03413.1"
/db_xref="GI:20978308"
/translation="MALPYHIFLFTVLLPSFTLLAPPCRCMTSSSPYOEFLMORP
GNIDAPYSRSLSKGTPPTATHTHPRNCYHSATLCMAHTHYTGKMINPSCPGGLV
TVCWTYFTQGMDSGGVODARREKHVEVLSQTRVHGTSPPYKGLDLSLHETLRT
HTRLVSLFNTLLGLHEVSAQNPNCWICLPINRPVYSIPVPOMNNFSEINTTSV
LVGLVSNLEITHRSNLTVCVFSNTYTTNSQICRWYTPPOIVCLPSGIFVCGTSA
YRCLNGSSEMCFLSPVPMPTIYTEDOLYVSKPRNKRPVILPVIAGVIGALG
TGIGITTSQFYKLSQELNGDMERVADSVTLQDOLNSIAAVLGNRRALDLITAE
RGGTCFLGEECCYVNOGSLVTEKVEIRRIORRAELRNTPGMGLISQMMWILP
FLGPLAIIILLFLGPCIFNLVNFVSSRIEAVKLQMPKQSTKTKYRRPLDRPAS
RSDVNDIKGTPPEIISAQPLRN"
BASE COUNT 449 a 456 c 290 g 429 t
ORIGIN

Query Match 100.0%; Score 201; DB 9; Length 1624;
Best Local Similarity 100.0%; Pred. No. 1e-48;
Matches 201; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AACTTCAGCAGCAAGAAATTAACACACCTCCGTTTAGTAGAGACCTCTGTTCCATCTG 60

```
|||||
Db 646 AACTTACGACAGCAAAATAAACACACCTCCGTTTGTAGACCTCTGTTCCAAATCTG 705
QY 61 GAAATAACCCATACCTCAAACTCAGCTGTGTAAATTTAGCAATCTACATACACAAAC 120
Db 706 GAAATAACCCATACCTCAAACTCAGCTGTGTAAATTTAGCAATCTACATACACAAAC 755
QY 121 AACTCCCAATGCATCAGGTGGGTAACTCTCCACACAAATAGTGTGCTACCTCAGGA 180
Db 766 AACTCCCAATGCATCAGGTGGGTAACTCTCCACACAAATAGTGTGCTACCTCAGGA 825
QY 181 ATATTTTGTCTGTGTAC 201
Db 826 ATATTTTGTCTGTGTAC 846

RESULT 11
AF513360 1860 bp mRNA linear PRI 05-JUN-2002
LOCUS Homo sapiens enverlin mRNA, complete cds.
DEFINITION AF513360
ACCESSION AF513360
VERSION AF513360.1 GI:21326140
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
REFERENCE 1 (bases 1 to 1860)
AUTHORS Alliel,P.M., Perl,J.P., Pleris,R., Nussbaum,J.L., Menard,A. and
Rieger,F.
TITLE Endogenous retroviruses and multiple sclerosis. Part 2: HERV-7q and
its env transcripts
C. R. Acad. Sci. III, Sect. Vie 312, 857-863 (1998)
2 (bases 1 to 1860)
AUTHORS Alliel,P.M., Perl,J.P., Goudou,D., Bitoun,M., Robert,B. and
Rieger,F.
TITLE The HERV-7q family in the human genome. Potential for protein
expression and gene regulation
Cell Mol. Biol. 48 (2), 213-217 (2002)
JOURNAL MEDLINE 21985840
PUBMED 11990458
REFERENCE 3 (bases 1 to 1860)
AUTHORS Alliel,P.M., Goudou,D., Perl,J.P. and Rieger,F.
TITLE Direct Submission
Submitted (18-MAY-2002) U-488, INSERM, 80, rue du General Leclerc,
Le Kremlin-Bicetre 94270, France
FEATURES
source
1. 1860
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="7"
/map="7q21"
/issue_type="Placenta"
59..1675
/codon_start=1
/product="enverlin"
/protein_id="AA047599.1"
/db_xref="GI:21326141"
/translation="MALPHTLFLTVLLPSTVLAPPCRCMTSSPYOELMRQRP
GNIDAPSYRLSKGPTPTAHTHMPNRCNCHSATLMAHTHTGKIMNPSCPGIGV
TYCMTYFTDTGMSDGGVODAREKHVKVILQLRVHSTSPYGLDLSKHEPLRT
HTRLVSLFTTLTGLHEVSAONPTNCMILPLNFPYVSIIPVEOMNNSFEINTSY
LVGPLVSNLEIHTSNLVCVFNSTNTYNSCIRNVPPIVCGISGIFVCGTSA
YKLGSSSESMCELSFLVAPMTITTEODLYNVISKPRKRVPIIPVIGAVIGAG
TGIGITTSSTORYKYSLOELNDEMRVDSITLDOQLNSLAAYLQNRALDLTAE
RGCTCLFLGEECCYVYNSGITYEKVEKTRDIOKRAEELRTGMLGSOWPILP
FLGPLAALLLFLGFCIPNLVNFVSSKIEAVKLOMEPKMOSKTKIYRPLDRPAS
RSDVNDIKTPEEISAPLIRPSAGSS"
BASE COUNT 513 a 520 c 350 g 477 t
ORIGIN
Query Match 100.0%; Score 201; DB 9; Length 1860;
Best Local Similarity 100.0%; Pred. No. 1e-48;
```

```
Matches 201; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 AACTTACGACAGCAAAATAAACACACCTCCGTTTGTAGACCTCTGTTCCAAATCTG 60
Db 680 AACTTACGACAGCAAAATAAACACACCTCCGTTTGTAGACCTCTGTTCCAAATCTG 739
QY 61 GAAATAACCCATACCTCAAACTCAGCTGTGTAAATTTAGCAATCTACATACACAAAC 120
Db 740 GAAATAACCCATACCTCAAACTCAGCTGTGTAAATTTAGCAATCTACATACACAAAC 799
QY 121 AACTCCCAATGCATCAGGTGGGTAACTCTCCACACAAATAGTGTGCTACCTCAGGA 180
Db 800 AACTCCCAATGCATCAGGTGGGTAACTCTCCACACAAATAGTGTGCTACCTCAGGA 859
QY 181 ATATTTTGTCTGTGTAC 201
Db 860 ATATTTTGTCTGTGTAC 880

RESULT 12
AX000960 1948 bp DNA linear PAT 10-MAR-2000
LOCUS AX000960
DEFINITION Sequence 5 from Patent WO9902656.
ACCESSION AX000960
VERSION AX000960.1 GI:7241202
KEYWORDS
SOURCE unidentified.
ORGANISM unidentified.
REFERENCE 1 (bases 1 to 1948)
AUTHORS Beseme,F. and Blond,J.
TITLE ENDOGENETIC RETROVIRAL SEQUENCES, ASSOCIATED WITH AUTOIMMUNE
DISEASES OR WITH PREGNANCY DISORDERS
Patent: WO 9902656-A 5 21-JAN-1999;
JOURNAL BIO MERIEUX (FR); BESEME FREDERIC (FR)
FEATURES
source
1. 1948
/organism="unidentified"
/db_xref="taxon:32644"
BASE COUNT 505 a 549 c 381 g 510 t 3 others
ORIGIN
Query Match 100.0%; Score 201; DB 6; Length 1948;
Best Local Similarity 100.0%; Pred. No. 1e-48;
Matches 201; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 AACTTACGACAGCAAAATAAACACACCTCCGTTTGTAGACCTCTGTTCCAAATCTG 60
Db 1340 AACTTACGACAGCAAAATAAACACACCTCCGTTTGTAGACCTCTGTTCCAAATCTG 1399
QY 61 GAAATAACCCATACCTCAAACTCAGCTGTGTAAATTTAGCAATCTACATACACAAAC 120
Db 1400 GAAATAACCCATACCTCAAACTCAGCTGTGTAAATTTAGCAATCTACATACACAAAC 1459
QY 121 AACTCCCAATGCATCAGGTGGGTAACTCTCCACACAAATAGTGTGCTACCTCAGGA 180
Db 1460 AACTCCCAATGCATCAGGTGGGTAACTCTCTCCACACAAATAGTGTGCTACCTCAGGA 1519
QY 181 ATATTTTGTCTGTGTAC 201
Db 1520 ATATTTTGTCTGTGTAC 1540

RESULT 13
AX027474 1948 bp DNA linear PAT 16-SEP-2000
LOCUS AX027474
DEFINITION Sequence 24 from Patent FR2788784.
ACCESSION AX027474
VERSION AX027474.1 GI:10188438
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
```


REFERENCE 1 (bases 1 to 1948)
AUTHORS Mallet, F., Voisset, C. and Paranhos, B. G.
JOURNAL Patent: FR 2788784-A 24 28-JUL-2000;
BIO MERIEUX (FR)

FEATURES
Source Location/Qualifiers
1..1948
/organism="Homo sapiens"
/db_xref="taxon:9606"

BASE COUNT 505 a 549 c 381 g 510 t 3 others

ORIGIN

Query Match 100.0%; Score 201; DB 6; Length 1948;
Best Local Similarity 100.0%; Pred. No. 1e-48;
Matches 201; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AACTTCAGCAGAAATAAACACCACTTCGCTTTTATAGAGACCTCTGTTTCCAACTCG 60
|||||
DB 1340 AACTTCAGCAGAAATAAACACCACTTCGCTTTTATAGAGACCTCTGTTTCCAACTCG 1399
|||||

QY 61 GAATAAACCCATACCTCAACCTGCTGTAATAATTAGCAATACATACACCAACC 120
|||||
DB 1400 GAATAAACCCATACCTCAACCTGCTGTAATAATTAGCAATACATACACCAACC 1459
|||||

QY 121 AACTCCCAATGATCAGGTGGGTAACTCTCCACACAAATAGTCTGCTACCTCAGGA 180
|||||
DB 1460 AACTCCCAATGATCAGGTGGGTAACTCTCCACACAAATAGTCTGCTACCTCAGGA 1519
|||||

QY 181 ATATTTTGTCTGTGTGTTACC 201
|||||
DB 1520 ATATTTTGTCTGTGTGTTACC 1540
|||||

RESULT 14
LOCUS AF072505 1948 bp mRNA linear PRI 10-FEB-1999
DEFINITION Homo sapiens endogenous retrovirus W envelope protein mRNA, partial
cgs.
ACCESSION AF072505
VERSION AF072505.1 GI:4262289
KEYWORDS
SOURCE Homo sapiens.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 1948)
AUTHORS Blond, J. L., Beseme, F., Duret, L., Bouton, O., Bedin, F., Perron, H.,
Mandrand, B. and Mallet, F.
TITLE Molecular characterization and placental expression of HERV-W, a
new human endogenous retrovirus family
JOURNAL J. Virol. 73 (2), 1175-1185 (1999)
MEDLINE 9882319
PUBMED 9882319

REFERENCE 2 (bases 1 to 1948)
AUTHORS Blond, J. L., Beseme, F. and Mallet, F.
TITLE Direct Submission
JOURNAL Submitted (19-JUN-1998) UM103 CNRS-Biomerieux, ENS Lyon, 46 allée
d'Italie, Lyon, Cedex 07 69364, France

FEATURES
Source Location/Qualifiers
1..1948
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="cl.24.4"
/tissue_type="placenta"
/clone_lib="Clontech 5'-stretch plus library, Cat number
HL5014a"
1..1948
/rpl_family="HERV-W"
<1..6
/note="R"
5'UTR 7..460
/note="U5"
primer_bind 464..481

misc_feature
491..492
/note="putative primer binding site for tRNA-W"

CDS
719..>1948
/codon_start=1
/product="envelope protein"
/protein_id="A01454.1"
/db_xref="GI:4262290"
/translation="MALPYHIFLFTVSPSELTLPAPPCRMNTSSSPYOEFLRMORP
GNDAPYSRISCKGPTPTFAHTHNPNCYHSATLGMHANTHMYTKMINSCEGIGV
IWCWTYFTOTGMSDGGVDOAKRKVKYISLTLYRHGSSPYKGLDLSKLEHETLT
HRLVSLNFTLTGLHVSANPTNCWICLPNRPVSIPEEOMNNESTELNTTSV
LVGPLVSNLEITHSNLTCVKSNTTYTNSQICIRWVTPPTQIVCLPSGIFVCGSA
YRCLNSSSESMCFLSFLVPMAYTEODLYSYISRPKRNPVILPEFVIGAGVLGALG
TGIGITTSYOFYKXLSOELNMGDMERAVSLVYLQQLNSLAIVLQNRALDLTAE
RGSTCLFLGECCXYVNO"

BASE COUNT 505 a 549 c 381 g 510 t 3 others

ORIGIN

Query Match 100.0%; Score 201; DB 9; Length 1948;
Best Local Similarity 100.0%; Pred. No. 1e-48;
Matches 201; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AACTTCAGCAGAAATAAACACCACTTCGCTTTTATAGAGACCTCTGTTTCCAACTCG 60
|||||
DB 1340 AACTTCAGCAGAAATAAACACCACTTCGCTTTTATAGAGACCTCTGTTTCCAACTCG 1399
|||||

QY 61 GAATAAACCCATACCTCAACCTGCTGTAATAATTAGCAATACATACACCAACC 120
|||||
DB 1400 GAATAAACCCATACCTCAACCTGCTGTAATAATTAGCAATACATACACCAACC 1459
|||||

QY 121 AACTCCCAATGATCAGGTGGGTAACTCTCCACACAAATAGTCTGCTACCTCAGGA 180
|||||
DB 1460 AACTCCCAATGATCAGGTGGGTAACTCTCCACACAAATAGTCTGCTACCTCAGGA 1519
|||||

QY 181 ATATTTTGTCTGTGTGTTACC 201
|||||
DB 1520 ATATTTTGTCTGTGTGTTACC 1540
|||||

RESULT 15
LOCUS AX007999 2055 bp DNA linear PAT 06-SEP-2000
DEFINITION Sequence 22 from Patent WO967395.
ACCESSION AX007999
VERSION AX007999.1 GI:9995696
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 2055)
AUTHORS Perin, J. P., Rieger, F. and Alliel, P. M.
TITLE Nucleic sequence and deduced protein sequence family with human
endogenous retroviral motifs, and their uses
JOURNAL INST NAT SANTE RECH MED (FR); PERIN JEAN PIERRE (FR); RIEGER
FRANCOIS (FR); ALLIEL PATRICK M (FR)
Location/Qualifiers
1..2055
/organism="Homo sapiens"
/db_xref="taxon:9606"
<1..87
/note="unnamed protein product"
/codon_start=1
/protein_id="CAC07392.1"
/db_xref="GI:9995697"
/translation="PKTANLVADITSXAKYQVYLKTLQGTYP"

BASE COUNT 576 a 574 c 376 g 529 t

ORIGIN

Query Match 100.0%; Score 201; DB 6; Length 2055;
Best Local Similarity 100.0%; Pred. No. 1e-48;

	Matches	201;	Conservative	0;	Mismatches	0;	Indels	0;	Gaps	0;
QY	1	AACTTCAGCAGCAAAATAAACAACCACTTCGCTTTAGTAGAGACCTCTGTGTTCCAAATCTG	60							
Db	1012	AACTTCAGCAGCAAAATAAACAACCACTTCGCTTTAGTAGAGACCTCTGTGTTCCAAATCTG	1071							
QY	61	GAAATTAACCCATACCTCAAACTCAGCTGTGTAAATTTAGCAATATACATACACACACC	120							
Db	1072	GAAATTAACCCATACCTCAAACTCAGCTGTGTAAATTTAGCAATATACATACACACACC	1131							
QY	121	AACTCCCAATGCATCAGGTGGGTAACTCTCCACACAAATAGTGTGCTACCTCAGGA	180							
Db	1132	AACTCCCAATGCATCAGGTGGGTAACTCTCCACACAAATAGTGTGCTACCTCAGGA	1191							
QY	181	ATATTTTGTCTGTGTGTTAC	201							
Db	1192	ATATTTTGTCTGTGTGTTAC	1212							

Search completed: May 2, 2003, 15:26:49
 Job time : 693 secs

GenCore version 5.1.5
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 2, 2003, 12:49:35 ; Search time 159.333 Seconds
(without alignments)
2840.910 Million cell updates/sec

Title: US-09-719-554-3_COPY_8500_8700

Perfect score: 201
Sequence: 1 aactcagacagacaataa.....tatttttctgtgttacc 201

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2185239 segs, 112599159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

N_Geneseq_101002:*

1: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT:*
2: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT:*
3: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT:*
4: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT:*
5: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT:*
6: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT:*
7: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT:*
8: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT:*
9: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT:*
10: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT:*
11: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT:*
12: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT:*
13: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT:*
14: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:*
15: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT:*
16: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT:*
17: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT:*
18: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT:*
19: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:*
20: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:*
21: /SID2/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:*
22: /SID2/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:*
23: /SID2/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:*
24: /SID2/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	201	100.0	1617	22	AAH20070
2	201	100.0	1948	20	AAH25659
3	201	100.0	1948	21	AAH59209
4	201	100.0	2055	21	ABN97948
5	201	100.0	2599	21	ABN97927
6	201	100.0	2781	22	AAH55630
7	201	100.0	2930	24	AAH24195
8	201	100.0	2946	20	AAH75526
9	201	100.0	2946	21	AAH59468

10	201	100.0	10499	21	ABN97929
11	201	100.0	56093	24	ABH61744
12	200.6	99.8	7582	21	AAH59215
13	199.4	99.2	2782	20	AAH25661
14	199.4	99.2	2782	21	AAH59211
15	199.4	99.2	2782	22	AAH20069
16	199.4	99.0	7582	20	AAH25665
17	193	96.0	1481	19	AAH43217
18	193	96.0	1481	20	AAH29703
19	193	96.0	1629	21	AAH96625
20	193	96.0	1799	21	ABN97931
21	193	96.0	2030	21	AAH63826
22	193	96.0	46340	21	ABN97978
23	191.4	95.2	902	23	AAH71722
24	191.4	95.2	1039	23	AAH84206
25	191.4	95.2	1341	23	AAH71719
26	191.4	95.2	1341	23	AAH84194
27	175.4	87.3	792	22	ABA50940
28	175.4	87.3	792	22	ABA68919
29	175.4	87.3	792	22	ABA35872
30	175.4	87.3	792	22	AAH17251
31	175.4	87.3	792	22	AAH43040
32	175.4	87.3	792	22	AAH23803
33	175.4	87.3	792	22	AAH49115
34	175.4	87.3	792	22	AAH09412
35	175.4	87.3	792	24	ABH17105
36	175.4	87.3	1894	22	ABA45822
37	175.4	87.3	1894	22	ABA56337
38	175.4	87.3	1894	22	ABA25978
39	175.4	87.3	1894	22	AAH04516
40	175.4	87.3	1894	22	AAH30018
41	175.4	87.3	1894	22	AAH14608
42	175.4	87.3	1894	22	AAH35960
43	175.4	87.3	1894	22	AAH04422
44	175.4	87.3	1894	24	ABH04589
45	175.4	87.3	2784	21	ABN97930

ALIGNMENTS

AAH20070	standard; DNA: 1617 BP.
AAH20070:	
08-AUG-2001	(first entry)
HERV-W envelope protein G coding sequence.	
Human endogenous retrovirus; HERV-W; chromosome 7; env protein; envelope protein; multiple sclerosis-related superantigen; vaccine; surface antigen; transmembrane; multiple sclerosis; neuroprotective; antisense-therapy; autoimmune disorder; ds.	
Human endogenous retrovirus.	
Key	Location/Qualifiers
CDS	1..1617
FT	/*tag= a
FT	/product= "HERV-W envelope protein G"
XX	WO200131021-A1.
XX	
XX	03-MAY-2001.
XX	
XX	30-OCT-2000; 2000WO-EP10659.
XX	
XX	28-OCT-1999; 99EP-0402690.
XX	
XX	(UYGE-) UNIV GENEVE.
XX	

Human retroviral s
Colon adenocarcinoma
Human endogenous r
Human endogenous r
5' non coding, 3'
HERV-W envelope pr
Complete human end
Multiple sclerosis
Clone C15 from MSR
DNA encoding an en
Human retroviral s
Nucleotide sequenc
Human retroviral s
DNA encoding novel
DNA encoding novel
DNA encoding novel
DNA encoding novel
Human breast cell
Human foetal liver
Probe #14338 for g
Human brain expres
Human bone marrow
Probe #13736 for g
Probe #17801 used
Probe #9403 used t
Human genome-deriv
Human breast cell
Human foetal liver
Probe #4444 for ge
Human brain expres
Human bone marrow
Probe #4541 for ge
Probe #4666 used t
Probe #4413 used t
Human genome-deriv
Human retroviral s

PI Conrad B, Mach B;
XX WPI: 2001-316336/33.
DR P-PSDB: AAB75138.
XX
PT New human retrovirus HERV-W ENV proteins/peptides having superantigen
PT activity useful for diagnosing and treating multiple sclerosis -
XX
PS Claim 13; Fig 10; 94pp; English.
XX
CC On the basis of the PBS-t-RNA motif used for the classification of human
CC endogenous retrovirus (HERVs) the full length endogenous provirus which
CC was located on the long arm of human chromosome 7 (9q21-22) has been
CC designated HERV-W. The present invention describes proteins or peptides
CC (I) having superantigen (Sag) activity comprising the ENV protein (ENV)
CC of HERV-W, the surface protein (SU) and transmembrane (TM) sub-units. (I)
CC have neuroprotective activity, and can be used in: vaccines; antisense-
CC therapy; and HERV-W Sag activity-inhibitors. (I) and encoding DNA/RNA are
CC useful for diagnosing multiple sclerosis (MS) or HERV-W-associated
CC disorders. (I) are also useful for identifying substances (and optionally
CC recovering) capable of binding to a retroviral superantigen associated
CC with MS, substances capable of blocking Sag activity and substances
CC capable of blocking transcription or translation of HERV-W retroviral
CC superantigen. A protein or peptide derived from (I), modified to be
CC devoid of Sag activity and being capable of generating an immune response
CC against HERV-W retroviral Sag is useful in therapy. Nucleic acid
CC molecules encoding (I) are useful as vaccines against MS. Substances
CC capable of blocking Sag activity, capable of binding to a retroviral
CC superantigen associated with MS, or capable of blocking transcription or
CC translation of HERV-W retroviral superantigen for use in treating or
CC prevention of MS. (I) and nucleic acids encoding them are useful for
CC diagnosing autoimmune disease. The present sequence encodes the
CC specifically claimed envelope protein of HERV-W designated G.
XX
SQ Sequence 1617 BP; 442 A; 452 C; 296 G; 427 T; 0 other;
XX
Query Match 100.0%; Score 201; DB 22; Length 1617;
Best Local Similarity 100.0%; Pred. No. 6.7e-52;
Matches 201; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
QY 1 AACTTCAGCAGAAATAAACAACCACTCCGTTTGTAGAGACCTCTGTTCCAAATC 60
DB 622 AACTTCAGCAGAAATAAACAACCACTCCGTTTGTAGAGACCTCTGTTCCAAATC 681
QY 61 GAATTAACCCATACCTCAAACTCAGCTGTGTAATTTAGCAATACATACACACAC 120
DB 682 GAATTAACCCATACCTCAAACTCAGCTGTGTAATTTAGCAATACATACACACAC 741
QY 121 AACTCCCATGATCAGTGGGTAACTCCTCCACACAAATAGTGGCTACCCCTAGGA 180
DB 742 AACTCCCATGATCAGTGGGTAACTCCTCCACACAAATAGTGGCTACCCCTAGGA 801
QY 181 ATATTTTGTCTGTGTGAC 201
DB 802 ATATTTTGTCTGTGTGAC 822
XX
RESULT 2
AAK25655
ID AAK25655 standard; CDNA to mRNA; 1948 BP.
XX
AC AAK25655;
XX
DT 21-MAY-1999 (first entry)
XX
XX Human endogenous retrovirus W clone cl.24.4.
XX
DE Human endogenous retrovirus W clone cl.24.4.
XX
KW Clone; human endogenous retrovirus; genome; autoimmune disease;
KW multiple sclerosis; rheumatoid polyarthritis; insulin-dependent diabetes;
KW disseminated lupus erythematosus; pregnancy; chromosomal marker; ss.
XX
OS Human endogenous retrovirus.

XX
PN WO9902696-A1.
XX
XX 21-JAN-1999.
PD
XX
XX 06-JUL-1998; 98MO-FR01442.
PF
XX 07-JUL-1997; 97FR-0008815.
PR
XX
PA (INMR) BIO MERIEUX.
PI Beseme F, Blond JL, Bouton O, Mallet F, Mandrand B;
XX WPI: 1999-120897/10.
XX
PT New nucleic acid sequences from human endogenous retrovirus-W -
PT expressed exclusively in placenta and useful in diagnosis and
PT therapy of autoimmune disease, and abnormal or failed pregnancy
XX
PS Claim 1; Page 56-58; 106pp; French.
XX
CC This sequence represents clone cl.24.4 of the human endogenous retrovirus
CC (HERV) W genome. The nucleic acids, their fragments or peptides encoded
CC by them are markers of autoimmune disease (e.g. multiple sclerosis,
CC rheumatoid polyarthritis, disseminated lupus erythematosus, insulin-
CC dependent diabetes and related pathologies) and of abnormal or
CC unsuccessful pregnancy and can be used as chromosomal markers for
CC susceptibility to these conditions, or proximity markers of genes
CC associated with this susceptibility.
XX
SQ Sequence 1948 BP; 505 A; 549 C; 381 G; 510 T; 3 other;
XX
Query Match 100.0%; Score 201; DB 20; Length 1948;
Best Local Similarity 100.0%; Pred. No. 7.1e-52;
Matches 201; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
QY 1 AACTTCAGCAGAAATAAACAACCACTCCGTTTGTAGAGACCTCTGTTCCAAATC 60
DB 1340 AACTTCAGCAGAAATAAACAACCACTCCGTTTGTAGAGACCTCTGTTCCAAATC 1399
QY 61 GAATTAACCCATACCTCAAACTCAGCTGTGTAATTTAGCAATACATACACACAC 120
DB 1400 GAATTAACCCATACCTCAAACTCAGCTGTGTAATTTAGCAATACATACACACAC 1459
QY 121 AACTCCCATGATCAGTGGGTAACTCCTCCACACAAATAGTGGCTACCCCTAGGA 180
DB 1460 AACTCCCATGATCAGTGGGTAACTCCTCCACACAAATAGTGGCTACCCCTAGGA 1519
QY 181 ATATTTTGTCTGTGTGAC 201
DB 1520 ATATTTTGTCTGTGTGAC 1540
XX
RESULT 3
AAA59209
ID AAA59209 standard; DNA; 1948 BP.
XX
AC AAA59209;
XX
DT 07-NOV-2000 (first entry)
XX
XX 5' non coding, 3' pol gene, and 5' env gene of HERV-W from human genome.
XX
KW Autoimmune disease; retrovirus; human endogenous retrovirus W; HERV-W;
KW gag gene; pregnancy; multiple sclerosis; T cell proliferation; ss.
XX
XX Homo sapiens.
OS
XX WO200043521-A2.
PN
XX 27-JUL-2000.
PD
XX 21-JAN-2000; 2000MO-FR00144.
PF

```
XX 21-JAN-1999; 99FR-0000888.
PR
XX
XX (INRM ) BIO MERIEUX.
PA
XX
XX Paranhos-Baccala G, Mallet F, Voliset C;
PI
XX WPI; 2000-499229/44.
DR
XX
XX New nucleic acid from human endogenous retrovirus, useful e.g. for
PT diagnosis of autoimmune disease and complications of pregnancy,
PT contains at least part of the gag gene
XX
XX PS Disclosure; Page 45; 53pp; French.
XX
XX The present sequence represents an endogenous retroviral nucleic acid
CC fragment, which is associated with an autoimmune disease, and is
CC integrated into the human genome. The fragment is originally derived
CC from a novel retrovirus, human endogenous retrovirus W (HERV-W). The
CC HERV-W retrovirus is associated with autoimmune disease, failure of
CC pregnancy or disorders of pregnancy. The nucleic acid fragment, or
CC proteins derived from it, are useful for diagnosis of autoimmune
CC disease (specifically multiple sclerosis) and for monitoring pregnancy.
CC The nucleic acid fragments may also be used for in situ labelling of
CC isolated chromosomes, while the transcription product can be used to
CC study or monitor T cell proliferation in vitro.
XX
XX SQ Sequence 1948 BP; 505 A; 549 C; 381 G; 510 T; 3 other;
Query Match 100.0%; Score 201; DB 21; Length 1948;
Best Local Similarity 100.0%; Pred. No. 7.1e-52;
Matches 201; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 AACTTCAGCAGCAAGAAATTAACACCACTTCCCTTTAGTAGACCTCTGTTCCAATCTG 60
DB 1340 AACTTCAGCAGCAAGAAATTAACACCACTTCCCTTTAGTAGACCTCTGTTCCAATCTG 1399
QY 61 GAAATTAACCCATACCTCAAACTCACTGCTGTAAATTTAGCAATACATACACCAACC 120
DB 1400 GAAATTAACCCATACCTCAAACTCACTGCTGTAAATTTAGCAATACATACACCAACC 1459
QY 121 AACTCCCAATGATCAGGTGGGTAACTCTCCACACAATAGTCTGCTACCTCAGAGA 180
DB 1460 AACTCCCAATGATCAGGTGGGTAACTCTCCACACAATAGTCTGCTACCTCAGAGA 1519
QY 181 ATATTTTGTCTGTGTGAC 201
DB 1520 ATATTTTGTCTGTGTGAC 1540
RESULT 4
ABN97948
ID ABN97948 standard; DNA; 2055 BP.
XX
XX AC ABN97948;
XX
XX 01-AUG-2002 (first entry)
XX
XX Human retroviral HERV-7g env coding sequence.
DE
XX
XX Autoimmune disease; HERV-7g; chromosome 7q; immunotherapy;
KW multiple sclerosis; ds.
XX
XX OS Human retrovirus.
XX
XX WO9967395-A1.
XX
XX 29-DEC-1999.
XX
XX 23-JUN-1999; 99WO-FR01513.
XX
XX 23-JUN-1998; 98FR-0007920.
```

```
PA (INRM ) INSERM INST NAT SANTE & RECH MEDICALE.
XX
XX Alliel PM, Perin J, Rieger F;
PI
XX WPI; 2000-160587/14.
DR
XX
XX New nucleic acid sequences of human endogenous retrovirus, HERV-7g,
PT used for diagnosis, treatment and prevention of autoimmune and
PT neurological diseases
XX
XX PS Claim 3; Page 142-145; 225pp; French.
XX
XX The present invention relates to new nucleic acid sequences of human
CC endogenous retrovirus, HERV-7g, which is located on chromosome 7q.
CC Regulatory elements associated with HERV-7g may alter expression of other
CC genes (even remote genes) on the same chromosome, inducing immunological
CC and/or neurological changes (which may be pathological or protective/
CC curative). HERV-7g peptides can be used to improve efficiency of the
CC immune response, e.g. in immunotherapy. HERV-7g peptides and their coding
CC sequences can be used in immunogenic or vaccinating compositions, for
CC protection against autoimmune diseases, particularly multiple sclerosis.
CC The peptides may also be used (by sequence comparison) to detect/identify
CC endogenous retroviruses that are abnormally expressed in cancer,
CC neuropathologies or other autoimmune diseases. The present sequence was
CC used to illustrate the invention.
XX
XX SQ Sequence 2055 BP; 576 A; 574 C; 376 G; 529 T; 0 other;
Query Match 100.0%; Score 201; DB 21; Length 2055;
Best Local Similarity 100.0%; Pred. No. 7.2e-52;
Matches 201; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 AACTTCAGCAGCAAGAAATTAACACCACTTCCCTTTAGTAGACCTCTGTTCCAATCTG 60
DB 1012 AACTTCAGCAGCAAGAAATTAACACCACTTCCCTTTAGTAGACCTCTGTTCCAATCTG 1071
QY 61 GAAATTAACCCATACCTCAAACTCACTGCTGTAAATTTAGCAATACATACACCAACC 120
DB 1072 GAAATTAACCCATACCTCAAACTCACTGCTGTAAATTTAGCAATACATACACCAACC 1131
QY 121 AACTCCCAATGATCAGGTGGGTAACTCTCCACACAATAGTCTGCTACCTCAGAGA 180
DB 1132 AACTCCCAATGATCAGGTGGGTAACTCTCCACACAATAGTCTGCTACCTCAGAGA 1191
QY 181 ATATTTTGTCTGTGTGAC 201
DB 1192 ATATTTTGTCTGTGTGAC 1212
RESULT 5
ABN97927
ID ABN97927 standard; DNA; 2599 BP.
XX
XX AC ABN97927;
XX
XX 01-AUG-2002 (first entry)
XX
XX Human retroviral sequence HERV 7 env.
DE
XX
XX Autoimmune disease; HERV-7g; chromosome 7q; immunotherapy;
KW multiple sclerosis; ds.
XX
XX OS Human retrovirus.
XX
XX WO9967395-A1.
XX
XX 29-DEC-1999.
XX
XX 23-JUN-1999; 99WO-FR01513.
XX
XX 23-JUN-1998; 98FR-0007920.
XX
XX (INRM ) INSERM INST NAT SANTE & RECH MEDICALE.
```

```
XX Allel PM, Perin J, Rieger F;
XX WPI: 2000-160587/14.
XX
XX New nucleic acid sequences of human endogenous retrovirus, HERV-7q,
XX used for diagnosis, treatment and prevention of autoimmune and
XX neurological diseases -
XX
XX Claim 1: Page 128-129; 225pp; French.
XX
XX The present invention relates to new nucleic acid sequences of human
XX endogenous retrovirus, HERV-7q, which is located on chromosome 7q.
XX Regulatory elements associated with HERV-7q may alter expression of other
XX genes (even remote genes) on the same chromosome, inducing immunological
XX and/or neurological changes (which may be pathological or protective/
XX curative). HERV-7q peptides can be used to improve efficiency of the
XX immune response, e.g. in immunotherapy. HERV-7q peptides and their coding
XX sequences can be used in immunogenic or vaccinating compositions, for
XX protection against autoimmune diseases, particularly multiple sclerosis.
XX The peptides may also be used (by sequence comparison) to detect/identify
XX endogenous retroviruses that are abnormally expressed in cancer,
XX neurophthologies or other autoimmune diseases. The present sequence was
XX used to illustrate the invention.
XX
XX Sequence 2599 BP; 744 A; 718 C; 495 G; 642 T; 0 other;
SQ
Query Match 100.0%; Score 201; DB 21; Length 2599;
Best Local Similarity 100.0%; Pred. No. 7.7e-52;
Matches 201; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 AACCTGACGACAGAAATAAACACCACTCCGTTTGTAGTAGGACCTCTGTTTCCAAATCG 60
DB 1550 AACCTGACGACAGAAATAAACACCACTCCGTTTGTAGTAGGACCTCTGTTTCCAAATCG 1609
QY 61 GAAATACCCATACCTCAACCTCAGCTGTTAAATTAGCAATATCTACTACATACACACC 120
DB 1610 GAAATACCCATACCTCAACCTCAGCTGTTAAATTAGCAATATCTACTACATACACACC 1669
QY 121 AACCTCCAAATGATCAGTGGTGAATCTCCACACAATAGTCTGCTACCCCTCAGGA 180
DB 1670 AACCTCCAAATGATCAGTGGTGAATCTCTCCACACAATAGTCTGCTACCCCTCAGGA 1729
QY 181 ATATTTTGTCTGTGTGTTACC 201
DB 1730 ATATTTTGTCTGTGTGTTACC 1750
RESULT 6
AAF55630
ID AAF55630 standard; DNA; 2781 BP.
XX
XX AAF55630;
XX
XX 29-MAY-2001 (first entry)
XX
XX Nucleotide sequence of a human endogenous retrovirus envelope protein.
XX
XX Envelope protein; HERV; syncytia formation; placental development;
XX syncytia; cancer; cell adhesion; ss.
XX
XX Human endogenous retrovirus.
XX
XX Key Location/Qualifiers
XX CDS 762..2378
XX FT /*tag= a
XX FT /product= "envelope protein"
XX
XX WO200116171-A1.
XX
XX 08-MAR-2001.
XX
XX 01-SEP-2000; 2000MO-FR02429.
```

```
XX 01-SEP-1999; 99FR-0011141.
XX 15-SEP-1999; 99FR-0011793.
XX
XX (INMR ) BIO MERIEUX.
XX (INRM ) INST NAT SANTE & RECH MEDICALE.
XX
XX Mallet F, Cosset F, Blond J, Lavillette D, Boulton O, Ruggieri A;
XX WPI: 2001-226676/23.
XX
XX P-PsDB: AAB67652.
XX
XX Detecting expression of human endogenous retrovirus envelope protein in
XX cells of a tissue or culture, from its ability to induce syncytia -
XX
XX Disclosure: Page 44-45; 57pp; French.
XX
XX The present sequence encodes a human endogenous retrovirus envelope
XX protein. The specification describes a method for detecting expression
XX of an envelope protein from a human endogenous retrovirus (HERV), in
XX cells, of a tissue or culture. The method comprises detecting syncytia
XX formation due to the fusogenic properties of the envelope protein.
XX Envelope polypeptides and polynucleotides are used to produce
XX therapeutic or prophylactic compositions, particularly for treatment of
XX cancer, to correct defects in placental development (or other natural
XX formation of other types of syncytia), and to promote adhesion of cells
XX in grafts or cellular repair processes. Expression of sequences
XX antisense to the polynucleotide are used to prevent formation of
XX syncytia.
XX
XX Sequence 2781 BP; 747 A; 770 C; 556 G; 708 T; 0 other;
SQ
Query Match 100.0%; Score 201; DB 22; Length 2781;
Best Local Similarity 100.0%; Pred. No. 7.8e-52;
Matches 201; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 AACCTGACGACAGAAATAAACACCACTCCGTTTGTAGTAGGACCTCTGTTTCCAAATCG 60
DB 1383 AACCTGACGACAGAAATAAACACCACTCCGTTTGTAGTAGGACCTCTGTTTCCAAATCG 1442
QY 61 GAAATACCCATACCTCAACCTCAGCTGTTAAATTAGCAATATCTACTACATACACACC 120
DB 1443 GAAATACCCATACCTCAACCTCAGCTGTTAAATTAGCAATATCTACTACATACACACC 1502
QY 121 AACCTCCAAATGATCAGTGGTGAATCTCCACACAATAGTCTGCTACCCCTCAGGA 180
DB 1503 AACCTCCAAATGATCAGTGGTGAATCTCTCCACACAATAGTCTGCTACCCCTCAGGA 1562
QY 181 ATATTTTGTCTGTGTGTTACC 201
DB 1563 ATATTTTGTCTGTGTGTTACC 1583
RESULT 7
AAD24195
ID AAD24195 standard; CDNA; 2930 BP.
XX
XX AAD24195;
XX
XX 07-MAY-2002 (first entry)
XX
XX Human syncytin cDNA.
XX
XX Human; syncytin; preclampsia; gestational trophoblast disorder;
XX choriocarcinoma; hydatiform mole; placental site tumour; abortion;
XX envelope gene; human endogenous defective retrovirus; HERV-W, ss.
XX
XX Homo sapiens.
XX
XX Key Location/Qualifiers
XX CDS 930..2546
XX FT /*tag= a
XX FT /product= "Syncytin"
```

XX MO200204678-A2.
PN 17-JAN-2002.
XX 09-JUL-2001; 2001WO-US21719.
XX 07-JUL-2000; 2000US-216657P.
XX (GEMV) GENETICS INST INC.
XX Keith JC, McCoy JM, Mi S;
XX PI
XX WPI: 2002-171727/22.
XX P-PSDB; AAEL4540.
XX Identifying a compound for treating a subject with or at risk of
XX developing preeclampsia, comprises determining whether the expression
XX or activity of syncytin in the cell is modulated in the presence of a
XX test compound
XX
XX Disclosure: Page 39-42; 43pp; English.
XX
XX The invention relates to identifying compounds which are modulators
XX of syncytin expression. The syncytin modulators are useful in diagnosis
XX and treatment of preeclampsia and gestational trophoblast disorders (e.g.
XX choriochorionoma, hydatiform mole, placental site tumour and missed/
XX incomplete abortion). Syncytin is a human gene derived from the
XX envelope gene of human endogenous defective retrovirus, HERV-W. The
XX present invention is based partly on the discovery that syncytin
XX expression is dramatically reduced in preeclampsia, and is also
XX mis-localised to the apical syncytiotrophoblast membrane. The present
XX sequence is human syncytin cDNA.
XX
SQ Sequence 2930 BP; 842 A; 800 C; 571 G; 717 T; 0 other;
Query Match 100.0%; Score 201; DB 24; Length 2930;
Best Local Similarity 100.0%; Pred. No. 7.9e-52;
Matches 201; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 AACTTCAGCAGAGAAATAAACACCACTTCGTTTATAGAGACCTCTGTTCCAAATCTG 60
DB 1551 AACTTCAGCAGAGAAATAAACACCACTTCGTTTATAGAGACCTCTGTTCCAAATCTG 1610
QY 61 GAATAATACCATACCTCAAACTCAGCTGTGTAAATTTAGCAATATATACACCAACC 120
DB 1611 GAATAATACCATACCTCAAACTCAGCTGTGTAAATTTAGCAATATATACACCAACC 1670
QY 121 AACTCCCAATGCATCAGGTGGGTAACTCTCCACACAAATAGTCTGCTACCCCTCAGGA 180
DB 1671 AACTCCCAATGCATCAGGTGGGTAACTCTCCACACAAATAGTCTGCTACCCCTCAGGA 1730
QY 181 ATATTTTGTGCTGTGGTACC 201
DB 1731 ATATTTTGTGCTGTGGTACC 1751
RESULT 8
AAAT7526
ID AAX77526 standard; cDNA; 2946 BP.
XX AC AAX77526;
XX 10-AUG-1999 (first entry)
XX DE Human secreted protein AJ172_2 cDNA.
XX
XX Secreted protein; testes; brain; blood; placenta; human; murine; thymus;
XX bone marrow; treatment; prevention; nutrition; cytokine; immune; vaccine;
XX cell proliferation; cell differentiation; suppressor; tumour inhibition;
XX haematopoiesis regulation; activin; inhibin; chemotactic; chemokine;
XX haemostatic; thrombolytic; receptor; ligand; anti-inflammatory; tumour;
XX cadherin; tumour invasion suppressor; gene therapy; tissue growth; ss.

XX OS Homo sapiens.
XX PN WO9926972-A1.
XX 03-JUN-1999.
XX
XX 17-NOV-1998; 98NO-US24614.
XX
XX 20-OCT-1998; 98US-0175928.
XX 21-NOV-1997; 97US-0976110.
XX 18-MAY-1998; 98US-0080478.
XX
XX (GEMV) GENETICS INST INC.
XX Collins-Racie LA, Evans C, Jacobs K, Lavallie ER;
XX McCoy JM, Merberg D, Treacy M;
XX WPI: 1999-357813/30.
XX P-PSDB; AAY08622.
XX
XX New polynucleotides encoding secreted proteins
XX
XX Claim 13a: Page 100-101; 142pp; English.
XX
XX This invention describes novel human secreted proteins encoded by
XX polynucleotides isolated from human adult testes, adult brain, adult
XX blood or adult placenta, or murine adult bone marrow or thymus cDNA
XX libraries. The products of the invention are predicted to have biological
XX activities which would make them suitable for treating, preventing or
XX ameliorating medical conditions in humans and animals, although no
XX supporting data is given. Suggested activities include nutritional
XX activity, cytokine and cell proliferation/differentiation activity,
XX immune stimulating (e.g. as vaccines) or suppressing activity,
XX haematopoiesis regulating activity, tissue growth activity,
XX activin/inhibin activity, chemotactic/chemokine activity, haemostatic
XX and thrombolytic activity, receptor/ligand activity, anti-inflammatory
XX activity, cadherin/tumour invasion suppressor activity, and tumour
XX inhibition activity. The polynucleotides are also stated to be useful
XX for gene therapy.
XX
SQ Sequence 2946 BP; 858 A; 801 C; 570 G; 717 T; 0 other;
Query Match 100.0%; Score 201; DB 20; Length 2946;
Best Local Similarity 100.0%; Pred. No. 8e-52;
Matches 201; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 AACTTCAGCAGAGAAATAAACACCACTTCGTTTATAGAGACCTCTGTTCCAAATCTG 60
DB 1549 AACTTCAGCAGAGAAATAAACACCACTTCGTTTATAGAGACCTCTGTTCCAAATCTG 1608
QY 61 GAATAATACCATACCTCAAACTCAGCTGTGTAAATTTAGCAATATATACACCAACC 120
DB 1609 GAATAATACCATACCTCAAACTCAGCTGTGTAAATTTAGCAATATATACACCAACC 1668
QY 121 AACTCCCAATGCATCAGGTGGGTAACTCTCCACACAAATAGTCTGCTACCCCTCAGGA 180
DB 1669 AACTCCCAATGCATCAGGTGGGTAACTCTCCACACAAATAGTCTGCTACCCCTCAGGA 1728
QY 181 ATATTTTGTGCTGTGGTACC 201
DB 1729 ATATTTTGTGCTGTGGTACC 1749
RESULT 9
AAZ59468
ID AAZ59468 standard; cDNA; 2946 BP.
XX AC AAZ59468;
XX 11-APR-2000 (first entry)
XX DE Human secreted protein AJ172_2 polynucleotide sequence.

XX Human; secreted protein; disease diagnosis; pre-eclampsia; cancer;
KW placental pathology; metastasis inhibition; nutritional activity;
KW immune stimulator; hematopoiesis regulator; tissue growth;
KW tumour inhibitor; anti-inflammatory; clone AJ172_2; ATCC_98115;
KW gene therapy; ss.
XX
OS Homo sapiens.
XX
PN WO960020-A1.
XX
PD 25-NOV-1999.
XX
PF 17-MAY-1999; 99WO-US10915.
XX
PR 18-MAY-1998; 98US-0080478.
PR 20-OCT-1998; 98US-0175928.
XX
PA (GENE) GENETICS INST INC.
XX
PI Jacobs K, McCoy JM, LaVallie ER, Collins-Racie LA, Evans C;
PI Merberg D, Ml S, Treacy M;
XX
DR MPI: 2000-116311/10.
DR P-PSDB: AAI67313.
XX
PT New polynucleotides encoding secreted cDNA libraries, used to develop
PT products for the diagnosis and treatment of neoplastic disease
XX
PS Claim 14; Page 107-108; 149pp: English.
XX
XX This is the human secreted protein AJ172_2 nucleotide sequence, obtained
CC from a human adult testes cDNA library. The invention relates to secreted
CC human and murine proteins. The polynucleotides and proteins are predicted
CC to have biological activities which would make them suitable for
CC treating, preventing or ameliorating medical conditions in humans and
CC animals. Detection of the levels of the proteins can be used for the
CC diagnosis of e.g. pre-eclampsia, placental pathology or cancer. Agents
CC which modulate the expression or function of the proteins may be used for
CC treating a neoplastic disease and inhibiting metastasis. Other suggested
CC activities include nutritional activity (e.g. in feeds), cytokine and
CC cell proliferation/differentiation activity, immune stimulating
CC (e.g. as vaccines) or suppressing activity, hematopoiesis regulating
CC activity, tissue growth activity, activin/inhibin activity,
CC chemotactic/chemokinetic activity, haemostatic and thrombolytic activity,
CC receptor/ligand activity, anti-inflammatory activity, cadherin/tumour
CC invasion suppressor activity, and tumour inhibition activity. The
CC polynucleotide sequences are also stated to be useful for gene therapy.
XX
SQ Sequence 2946 BP; 858 A; 801 C; 570 G; 717 T; 0 other;

Query Match 100.0%; Score 201; DB 21; Length 2946;
Best Local Similarity 100.0%; Pred. No. 8e-52;
Matches 201; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AACTTCAGCAGCAAGAAATAAACACCACTCCGTTTGTAGTAGAGACCTCTTGTTCAAATCTG 60
DB 1549 AACTTCAGCAGCAGAAATAAACACCACTCCGTTTGTAGTAGAGACCTCTTGTTCAAATCTG 1608
QY 61 GAAATTAACCCATACCTCAACCTCAACCTGTTAAATTTAGCAATTAATCTACATACACAAC 120
DB 1609 GAAATTAACCCATACCTCAACCTCAACCTGTTAAATTTAGCAATTAATCTACATACACAAC 1668
QY 121 AACTCCCAATGCATGAGTGGGTAACTCTCCACACAATAATAGTGTGCTTACCCTCAGAGA 180
DB 1669 AACTCCCAATGCATGAGTGGGTAACTCTCTCCACACAATAATAGTGTGCTTACCCTCAGAGA 1728
QY 181 AATATTTTGTCTGTGTGAC 201
DB 1729 AATATTTTGTCTGTGTGAC 1749

RESULT 10

ABN97929
ID ABN97929 standard; DNA; 10499 BP.
XX
AC ABN97929;
XX
DT 01-AUG-2002 (first entry)
XX
DE Human retroviral sequence HERV-7q.
XX
KW Autoimmune disease; HERV-7q; chromosome 7q; immunotherapy;
KW multiple sclerosis; ds.
XX
OS Human retrovirus.
XX
PN WO967395-A1.
XX
PD 29-DEC-1999.
XX
PF 23-JUN-1999; 99WO-FR01513.
XX
PR 23-JUN-1998; 98FR-0007920.
XX
PA (INRM) INSERM INST NAT SANTE & RECH MEDICALE.
XX
PI Alliel PM, Perin J, Rieger F;
XX
DR MPI: 2000-160587/14.
XX
PT New nucleic acid sequences of human endogenous retrovirus, HERV-7q,
PT used for diagnosis, treatment and prevention of autoimmune and
PT neurological diseases
XX
PS Claim 3; Fig 1; 225pp: French.
XX
XX The present invention relates to new nucleic acid sequences of human
CC endogenous retrovirus, HERV-7q, which is located on chromosome 7q.
CC Regulatory elements associated with HERV-7q may alter expression of other
CC genes (even remote genes) on the same chromosome, inducing immunological
CC and/or neurological changes (which may be pathological or protective/
CC curative). HERV-7q peptides can be used to improve efficiency of the
CC immune response, e.g. in immunotherapy. HERV-7q peptides and their coding
CC sequences can be used in immunogenic or vaccinating compositions, for
CC protection against autoimmune diseases, particularly multiple sclerosis.
CC The peptides may also be used (by sequence comparison) to detect/identify
CC endogenous retroviruses that are abnormally expressed in cancer,
CC neuropathologies or other autoimmune diseases. The present sequence was
CC used to illustrate the invention.
XX
SQ Sequence 10499 BP; 3048 A; 2676 C; 2280 G; 2495 T; 0 other;

Query Match 100.0%; Score 201; DB 21; Length 10499;
Best Local Similarity 100.0%; Pred. No. 1.1e-51;
Matches 201; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AACTTCAGCAGCAAGAAATAAACACCACTCCGTTTGTAGTAGAGACCTCTTGTTCAAATCTG 60
DB 8500 AACTTCAGCAGCAGAAATAAACACCACTCCGTTTGTAGTAGAGACCTCTTGTTCAAATCTG 8559
QY 61 GAAATTAACCCATACCTCAACCTCAACCTGTTAAATTTAGCAATTAATCTACATACACAAC 120
DB 8560 GAAATTAACCCATACCTCAACCTCAACCTGTTAAATTTAGCAATTAATCTACATACACAAC 8619
QY 121 AACTCCCAATGCATGAGTGGGTAACTCTCCACACAATAATAGTGTGCTTACCCTCAGAGA 180
DB 8620 AACTCCCAATGCATGAGTGGGTAACTCTCTCCACACAATAATAGTGTGCTTACCCTCAGAGA 8679
QY 181 AATATTTTGTCTGTGTGAC 201
DB 8680 AATATTTTGTCTGTGTGAC 8700

RESULT 11
ABL61744

PT	determining a change in expression of a gene of a signature gene set -
XX	
PS	Claim 1; SEQ ID 81; 44pp; English.
CC	The present invention describes a method (M1) for screening for an
CC	anti-neoplastic agent. The method involves exposing cells to a chemical
CC	agent to be tested for anti-neoplastic activity, determining a change in
CC	expression of at least one gene (I) of a signature gene set, where (I)
CC	comprises a sequence (S) selected from 8447 sequences (given in AB161664
CC	to AB170110), or is at least 95% identical to (S), where a change in
CC	expression is indicative of anti-neoplastic activity. (I) has cytostatic
CC	activity and can be used in gene therapy. M1 can be used for screening
CC	an anti-neoplastic agent, and can be used for producing a product which
CC	is the data collected with respect to the anti-neoplastic agent as a
CC	result of M1, and the data is sufficient to convey the chemical
CC	structure and/or properties of the agent. M1 can be used in the
CC	treatment of cancer such as colon, breast, stomach, lung, thyroid,
CC	oesophageal, ovarian, kidney, prostate or pancreatic cancer,
CC	adenocarcinoma, carcinoma, clear cell cancer, infiltrating ductal cancer,
CC	infiltrating lobular cancer, squamous cell carcinoma, neuroendocrine
CC	carcinoma, papillary carcinoma and Wilms' tumour.
CC	
SQ	Sequence 56093 BP; 16164 A; 12346 C; 10702 G; 16881 T; 0 other;
	Query Match 100.0%; Score 201; DB 24; Length 56093;
	Best Local Similarity 100.0%; Pred. No. 1.8e-51;
	Matches 201; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY	1 AACCTGACACAGAAATAAACGCCACTTCGGTTTAGTGAAGACCTGTGCCAATCTG 60
Db	36500 AACCTGCACACAGAAATAAACGCCACTTCGGTTTAGTGAAGACCTGTGCCAATCTG 36559
QY	61 GAATTAACCCCATACTCAAACTCACCCTGTGTAATAATTAGCAATACTACATACACAACC 120
Db	36560 GAATTAACCCCATACTCAAACTCACCCTGTGTAATAATTAGCAATACTACATACACAACC 36619
QY	121 AACTCCCAATGCGATCAGGTGGGTAACTCCTCCACACAAATAGTGCTGCCCTACCCTCAGGA 180
Db	36620 AACTCCCAATGCGATCAGGTGGGTAACTCCTCCACACAAATAGTGCTGCCCTACCCTCAGGA 36679
QY	181 ATATTTTTTGTCTGTGTGTACC 201
Db	36680 ATATTTTTTGTCTGTGTGTACC 36700
RESULT 12	
ID	AAA59215 standard; DNA: 7582 BP.
AC	AAA59215;
XX	
DT	07-NOV-2000 (first entry)
XX	
DE	Human endogenous retrovirus W (HERV-W) sequence.
XX	
KW	Autoimmune disease; retrovirus; human endogenous retrovirus W; HERV-W;
XN	gag gene; pregnancy; multiple sclerosis; T cell proliferation; ss.
OS	Human endogenous retrovirus.
XX	
FT	Key Location/Qualifiers
FT	LTR 1..120
FT	/tag= a
FT	/note= "R of 5' LTR"
FT	LTR 121..575
FT	/tag= b
FT	/note= "U5 of 5' LTR"
FT	/tag= c
FT	/tag= d
FT	primer_bind 5581..7194
FT	/tag= ORF1 env538"
FT	/note= "ORF1 env538"
FT	CDS 7039..7194
FT	

```
FT      /*tag= e
FT      /note= "ORF2 52 AA"
FT      7112..7255
FT      CDS      /*tag= f
FT      /note= "ORF3 48 AA"
FT      7244..7254
FT      misc_feature /*tag= g
FT      /note= "polypurine tract"
FT      7256..7582
FT      LTR      /*tag= h
FT      /note= "U3-R of 3' LTR"
FT      7563..7569
FT      polyA_signal /*tag= i
FT      1
FT
XX      WO200043521-A2.
XX
XX      27-JUL-2000.
XX
XX      21-JAN-2000; 2000WO-FR00144.
XX
XX      21-JAN-1999; 99FR-0000888.
XX
XX      (INMR ) BIO MERIEUX.
XX
XX      Paranhos-Baccala G, Mallet F, Voisset C;
XX      WPI; 2000-499229/44.
XX
XX      New nucleic acid from human endogenous retrovirus, useful e.g. for
XX      diagnosis of autoimmune disease and complications of pregnancy,
XX      contains at least part of the gag gene
XX
XX      Disclosure; Page 49-52; 53pp; French.
XX
XX      The present sequence represents an endogenous retrovirus, which is
XX      associated with an autoimmune disease, and is integrated into the human
XX      genome. The retrovirus is human endogenous retrovirus W (HERV-W). The
XX      HERV-W retrovirus is associated with autoimmune disease, failure of
XX      pregnancy or disorders of pregnancy. HERV-W nucleic acid fragments, or
XX      proteins derived from it, are useful for diagnosis of autoimmune
XX      disease (specifically multiple sclerosis) and for monitoring pregnancy.
XX      The nucleic acid fragments may also be used for in situ labelling of
XX      isolated chromosomes, while the transcription product can be used to
XX      study or monitor T cell proliferation in vitro.
XX
XX      Sequence 7582 BP; 2156 A; 1876 C; 1538 G; 1796 T; 216 other;
XX
XX      Query Match      99.8%; Score 200.6; DB 21; Length 7582;
XX      Best Local Similarity 99.5%; Pred. No. 1.4e-51;
XX      Matches 200; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
XX
XX      1 AACTTCAGCAGCAATAAATACACCACTCCGTTTAGTAGACCTCTGTTCCAATCG 60
XX      |||||||
XX      6202 AACTTCAGCAGCAATAAATACACCACTCCGTTTAGTAGAGACCTCTGTTCCAATG 6261
XX      |||||||
XX      61 GAAATTAACCCATACCTCAAACTCAGCTGTGTAATATTAGCAATACATACACAAACC 120
XX      |||||||
XX      6262 GAAATTAACCCATACCTCAAACTCAGCTGTGTAATATTAGCAATACATACACAAACC 6321
XX      |||||||
XX      121 AACTCCCAATGCATCAGGTGGGTAACCTCCACACAAATAGTCTGCTACCTCAGAGA 180
XX      |||||||
XX      6322 AACTCCCAATGCATCAGGTGGGTAACCTCCACACAAATAGTCTGCTACCTCAGAGA 6381
XX      |||||||
XX      181 ATATTTTGTCTGTGTGTTAC 201
XX      |||||||
XX      6382 ATATTTTGTCTGTGTGTTAC 6402
XX
XX      RESULT 13
XX      AAX25661
XX      ID AAX25661 standard; CDNA to mRNA; 2782 BP.
XX      AAX25661;
XX
```

```
XX      21-MAY-1999 (first entry)
XX
XX      Human endogenous retrovirus W clone cl.FH74.
XX
XX      Clone; human endogenous retrovirus; genome; autoimmune disease;
XX      multiple sclerosis; rheumatoid polyarthritis; insulin-dependent diabetes;
XX      disseminated lupus erythematosus; pregnancy; chromosomal marker; ss.
XX
XX      Human endogenous retrovirus.
XX
XX      WO9902696-A1.
XX
XX      21-JAN-1999.
XX
XX      06-JUL-1998; 98WO-FR01442.
XX
XX      07-JUL-1997; 97FR-0008815.
XX
XX      (INMR ) BIO MERIEUX.
XX
XX      Beseme F, Blond JL, Bouton O, Mallet F, Mandrand B;
XX      WPI; 1999-120897/10.
XX
XX      New nucleic acid sequences from human endogenous retrovirus-W -
XX      expressed exclusively in placenta and useful in diagnosis and
XX      therapy of autoimmune disease, and abnormal or failed pregnancy
XX      Claim 1; Page 60-63; 106pp; French.
XX
XX      This sequence represents clone cl.PH74 of the human endogenous retrovirus
XX      (HERV) W genome. The nucleic acids, their fragments or peptides encoded
XX      by them are markers of autoimmune disease (e.g. multiple sclerosis,
XX      rheumatoid polyarthritis, disseminated lupus erythematosus, insulin-
XX      dependent diabetes and related pathologies) and of abnormal or
XX      unsuccessful pregnancy and can be used as chromosomal markers for
XX      susceptibility to these conditions, or proximity markers of genes
XX      associated with this susceptibility.
XX
XX      Sequence 2782 BP; 741 A; 767 C; 565 G; 709 T; 0 other;
XX
XX      Query Match      99.2%; Score 199.4; DB 20; Length 2782;
XX      Best Local Similarity 99.5%; Pred. No. 2.4e-51;
XX      Matches 200; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
XX
XX      1 AACTTCAGCAGCAATAAATACACCACTCCGTTTAGTAGACCTCTGTTCCAATCG 60
XX      |||||||
XX      1384 AACTTCAGCAGCAATAAATACACCACTCCGTTTAGTAGAGACCTCTGTTCCAATG 1443
XX      |||||||
XX      61 GAAATTAACCCATACCTCAAACTCAGCTGTGTAATATTAGCAATACATACACAAACC 120
XX      |||||||
XX      1444 GAAATTAACCCATACCTCAAACTCAGCTGTGTAATATTAGCAATACATACACAAACC 1503
XX      |||||||
XX      121 AACTCCCAATGCATCAGGTGGGTAACCTCCACACAAATAGTCTGCTACCTCAGAGA 180
XX      |||||||
XX      1504 AACTCCCAATGCATCAGGTGGGTAACCTCCACACAAATAGTCTGCTACCTCAGAGA 1563
XX      |||||||
XX      181 ATATTTTGTCTGTGTGTTAC 201
XX      |||||||
XX      1564 ATATTTTGTCTGTGTGTTAC 1584
XX
XX      RESULT 14
XX      AAA59211
XX      ID AAA59211 standard; DNA; 2782 BP.
XX      AAA59211;
XX
XX      07-NOV-2000 (first entry)
XX
XX      5' non coding, 3' pol gene, and 3' non coding sequences of HERV-W.
XX
```

[illegible]

XX	KM		antisense-therapy; autoimmune disorder; ds.
OS	XX	Human endogenous retrovirus.	
FT	XX		
FT	XX	Key	Location/Qualifiers
FT	XX	5'UTR	1..762
FT	XX	CDS	/*tag= a 763..2379
FT	XX		/*tag= b
FT	XX		/product= "HERV-W envelope protein G"
FT	XX		/transl_except= (pos:790..792,aa:Phe)
FT	XX		/transl_except= (pos:793..795,aa:Thr)
FT	XX		/transl_except= (pos:812..814,aa:Ileu)
FT	XX		/transl_except= (pos:818..820,aa:Ser)
FT	XX		/transl_except= (pos:862..864,aa:Tyr)
FT	XX		/transl_except= (pos:865..867,aa:Gln)
FT	XX		/transl_except= (pos:1174..1176,aa:Arg)
FT	XX		/transl_except= (pos:1441..1443,aa:Leu)
FT	XX		/transl_except= (pos:1903..1905,aa:Gln)
FT	XX		/transl_except= (pos:2017..2019,aa:Lys)
FT	XX		/transl_except= (pos:2026..2028,aa:Arg)
FT	XX		/transl_except= (pos:2044..2046,aa:Arg)
FT	XX		/transl_except= (pos:2089..2091,aa:Gln)
FT	XX		/transl_except= (pos:2170..2172,aa:Asn)
FT	XX	3' UTR	2380..2782
FT	XX		/*tag= c
PM	XX	WO200131021-A1.	
PD	XX	03-MAY-2001.	
XX	XX	30-OCT-2000; 2000WO-EPI0659.	
PR	XX	28-OCT-1999; 99EP-0402690.	
PA	XX	(UYGE-) UNIV GENEVE.	
PI	XX	Conrad B, Mach B;	
DR	XX	WPI; 2001-316336/33.	
XX	XX	P-PSDB; AAB75138.	
PT	XX	New human retrovirus HERV-W ENV proteins/peptides having superantigen	
PS	XX	activity useful for diagnosing and treating multiple sclerosis -	
XX	XX	Claim 13; Fig 9; 94pp; English.	
CC	XX	On the basis of the PBS t-RNA motif used for the classification of human	
CC	XX	endogenous retrovirus (HERVs) the full length endogenous provirus which	
CC	XX	was located on the long arm of human chromosome 7 (7q21-22) has been	
CC	XX	designated HERV-W. The present invention describes proteins or peptides	
CC	XX	(I) having superantigen (SAG) activity comprising the ENV protein (ENV)	
CC	XX	of HERV-W, the surface protein (SU) and transmembrane (TM) sub-units. (I)	
CC	XX	have neuroprotective activity, and can be used in vaccines; antisense-	
CC	XX	therapy; and HERV-W SAG activity-inhibitors. (I) and encoding DNA/RNA are	
CC	XX	useful for diagnosing multiple sclerosis (MS) or HERV-W-associated	
CC	XX	disorders. (I) are also useful for identifying substances (and optionally	
CC	XX	recovering) capable of binding to a retroviral superantigen associated	
CC	XX	with MS, substances capable of blocking SAG activity and substances	
CC	XX	capable of blocking transcription or translation of HERV-W retroviral	
CC	XX	superantigen. A protein or peptide derived from (I), modified to be	
CC	XX	devoid of SAG activity and being capable of generating an immune response	
CC	XX	against HERV-W retroviral SAG is useful in therapy. Nucleic acid	
CC	XX	molecules encoding (I) are useful as vaccines against MS. Substances	
CC	XX	capable of blocking SAG activity, capable of binding to a retroviral	
CC	XX	superantigen associated with MS, or capable of blocking transcription or	
CC	XX	translation of HERV-W retroviral superantigen for use in treating or	
CC	XX	preventing MS, obtained using (I) are useful for the treatment and	
CC	XX	prevention of MS. (I) and nucleic acids encoding them are useful for	
CC	XX	diagnosing autoimmune disease. The present sequence encodes the	
CC	XX	specifically claimed envelope protein of HERV-W designated G.	
SQ	XX	Sequence 2782 BP; 741 A; 768 C; 564 G; 709 T; 0 other:	

Query Match 99.2%; Score 199.4; DB 22; Length 2782;
Best Local Similarity 99.5%; Pred. No. 2.4e-51;
Matches 200; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AACTTCAGCAGAGAAATAAACACCACTTCGCTTTTAGTAGGACCTCTGTGTTCCAAATCTG 60
|||||
Db 1384 AACTTCAGCAGAGAAATAAACACCACTTCGCTTTTAGTAGGACCTCTGTGTTCCAAATGTG 1443
|||||

QY 61 GAATAAACCCTACCTCAACCTCACCTGTGTAAATTTAGCAATACATACACACACC 120
|||||
Db 1444 GAATAAACCCTACCTCAACCTCACCTGTGTAAATTTAGCAATACATACACACACC 1503
|||||

QY 121 AACTCCCAATGATCAGGTGGTAACTCTCCACACAAATAGTGTGCTACCCCTCAGGA 180
|||||
Db 1504 AACTCCCAATGATCAGGTGGTAACTCTCCACACAAATAGTGTGCTACCCCTCAGGA 1563
|||||

QY 181 ATATTTTTTGTCTGTGTACC 201
|||||
Db 1564 ATATTTTTTGTCTGTGTACC 1584
|||||

Search completed: May 2, 2003, 14:51:44
Job time : 184.333 secs

GenCore version 5.1.5
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 2, 2003, 14:19:30 ; Search time 1112 Seconds
(without alignments)
2927.422 Million cell updates/sec

Title: US-09-719-554-3_COPY_8500_8700

Perfect score: 201
Sequence: 1 aactcagcagcagaataaa.....tatttttgcgtgtacc 201

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hcc:*
9: gb_est1:*
10: gb_est2:*
11: gb_hcc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pln:*
21: em_gss_vrt:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_trod:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	201	100.0	459	14	R68740 y114g06.r1
2	201	100.0	723	9	AU138405 AU138405
3	193	96.0	631	17	AQ261133 C17B1-E1-
4	190	94.5	702	9	AU138097 AU138097
5	171.2	85.2	346	13	BI053386 RC5-GN028
6	171	85.1	315	13	BI055459 RC5-GN028

	C	7	171	85.1	315	13	BI056080
	C	8	171	85.1	376	14	BI056595
	C	9	170.8	85.0	374	14	T93661 y605g09.r1
	C	10	169.4	84.3	345	13	BI056592 RC5-GN028
	C	11	169	84.1	313	13	BI056596 RC5-GN028
	C	12	169	84.1	373	13	BI052559 RC5-GN028
	C	13	169	84.1	373	13	BI052567 RC5-GN028
	C	14	169	84.1	379	13	BI052566 RC5-GN028
	C	15	169	84.1	382	13	BI052565 RC5-GN028
	C	16	167.4	83.3	373	13	BI056589 RC5-GN028
	C	17	167.4	83.3	374	13	BI056593 RC5-GN028
	C	18	165.8	82.5	300	13	BI055456 RC5-GN028
	C	19	165.8	82.5	300	13	BI056077 RC5-GN028
	C	20	165.8	82.5	373	13	BI052561 RC5-GN028
	C	21	165.8	82.5	373	13	BI052571 RC5-GN028
	C	22	165.8	82.5	373	13	BI052572 RC5-GN028
	C	23	164.2	81.7	373	13	BI052570 RC5-GN028
	C	24	158.8	79.0	331	13	BI055458 RC5-GN028
	C	25	158.8	79.0	331	13	BI056079 RC5-GN028
	C	26	154.8	77.0	334	13	BI053385 RC5-GN028
	C	27	116.8	58.1	586	10	BE019603 ba84f03.y
	C	28	107	53.2	474	14	R00743 y674h09.r1
	C	29	101	50.2	267	14	H13097 y110h08.r1
	C	30	100	49.8	714	12	BE734284 601565487
	C	31	91	45.3	128	14	T39519 ya07b04.r1
	C	32	84	41.8	586	17	AQ776289 HS_5528.A
	C	33	81	40.3	790	12	BS572445 602593450
	C	34	79.6	39.6	718	9	AL601847 DKF2P313K
	C	35	69.4	34.5	317	13	BI056590 RC5-GN028
	C	36	39	19.4	349	14	BQ366047 RC6-GN007
	C	37	37.4	18.6	939	17	AL059400 Drosophila
	C	38	36.6	18.2	717	13	BM638420 170006875
	C	39	36.2	18.0	883	17	AZ671602 ENTIC17TF
	C	40	36.2	18.0	919	17	AQ749036 HS_5575.A
	C	41	36.2	18.0	922	17	A2686568 ENTKQ46TR
	C	42	35.8	17.8	732	17	AL099769 Drosophila
	C	43	35.6	17.7	430	9	AA058696 zf58B04.r
	C	44	35.6	17.7	1101	17	AL078714 Drosophila
	C	45	17.4	640	10	10	BB538097 BB538097

ALIGNMENTS

RESULT 1
R68740
LOCUS
DEFINITION
R68740 459 bp mRNA linear EST 26-JUN-1995
y114g06.r1 Soares placenta NB2HP Homo sapiens CDNA clone
IMAGE:139258 5' similar to SP:B44282 B44282 RETROVIRUS-RELATED ENV
POLYPROTEIN -, mRNA sequence.

ACCESSION
R68740
VERSION
R68740.1 GI:842257
KEYWORDS
EST.

SOURCE
ORGANISM
human.

REFERENCE

AUTHORS

TITLE
JOURNAL
COMMENT
The WashU-Merck EST Project
Unpublished (1995)
Contact: Wilson RK
Washington University School of Medicine
444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@wustl.edu
Insert Size: 1644
High quality sequence stops: 316
Source: IMAGE Consortium, LLNL


```
/clone_id="CITBf-EI"
/sex="male"
/cell_type="sperm"
/note="Vector: pBelobAC11; Site_1: EcoRI; Site_2: EcoRI;
Caltech Human BAC library D"
BASE COUNT      182 a      189 c      99 g      161 t
ORIGIN

Query Match      96.0%; Score 193; DB 17; Length 631;
Best Local Similarity 97.5%; Pred. No. 4.6e-50;
Matches 196; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 AACTTCAGCAGAAATAAACACCACTCCCTTTAGTAGAGACCTCTGTTCCATCTG 60
    |||||||
Db 245 AACTTCAGCAGAAATAAACACCACTCCCTTTAGTAGAGACCTCTGTTCCATCTG 304
    |||||||

QY 61 GAAATTAACCCATACCTCAACCTGCTGTAAATTTAGCAATACATACACACAC 120
    |||||||
Db 305 GAAATTAACCCATACCTCAACCTGCTGTAAATTTAGCAATACATACACACAC 364
    |||||||

QY 121 AACGCCAATGCATCAGGTGGTACCTCCACACAAATAGTCTGCTACCTCAGCA 180
    |||||||
Db 365 AGCTCCCAATGCATCAGGTGGTACCTCCACACAAATAGTCTGCTACCTCAGCA 424
    |||||||

QY 181 ATATTTTGTCTGTGTAC 201
    |||||||
Db 425 ATATTTTGTCTGTGTAC 445
    |||||||

RESULT 4
AUI38097      702 bp      mRNA      linear      EST 02-AUG-2002
LOCUS        AUI38097 PLACE1 Homo sapiens cDNA clone PLACE1007839 5', mRNA
DEFINITION   AUI38097 PLACE1 Homo sapiens cDNA clone PLACE1007839 5', mRNA
sequence.
ACCESSION    AUI38097
VERSION      AUI38097.1 GI:10999618
KEYWORDS     EST.
SOURCE       human.
ORGANISM     Homo sapiens
              Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
              Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE    1 (bases 1 to 702)
              Oca,T., Nishikawa,T., Suzuki,Y., Ishii,S., Saito,K., Kawai,Y.,
              Yamamoto,J., Wakamatsu,A., Nakamura,Y., Nagai,T., Sugano,S. and
              Isogai,T.
              HRI human cDNA project
              Unpublished (2000)
              Contact: Takao Isogai
              Genomics Laboratory
              Helix Research Institute
              1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
              Tel: 81-438-52-3975
              Fax: 81-438-52-3986
              Email: genomics@hri.co.jp
              HRI human cDNA project; 5'- & 3'-end one pass sequencing: Helix
              Research Institute; cDNA library construction: Department of
              Virology, Institute of Medical Science, University of Tokyo, and
              Helix Research Institute.
              Location/Qualifiers
                source
                  1. 702
                    /organism="Homo sapiens"
                    /db_xref="taxon:9606"
                    /clone_id="PLACE1007839"
                    /clone_lib="PLACE1"
                    /tissue_type="placenta"
                    /note="Vector: pME18SF13"
BASE COUNT      206 a      173 c      131 g      189 t      3 others
ORIGIN

Query Match      94.5%; Score 190; DB 9; Length 702;
Best Local Similarity 100.0%; Pred. No. 4.2e-49;
Matches 190; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 12 AGAATAAACAACACCTCCCTTTAGTAGAGACCTCTGTTCCATCTGAAATAACCA 71
    |||||||
Db 1 AGAATAAACAACACCTCCCTTTAGTAGAGACCTCTGTTCCATCTGAAATAACCA 60
    |||||||

QY 72 TACCTCAACCTCAACCTGTAAATTTAGCAATACATACACACCACTCCCATG 131
    |||||||
Db 61 TACCTCAACCTCAACCTGTAAATTTAGCAATACATACACACCACTCCCATG 120
    |||||||

QY 132 CATCAGGTGGTAACTCTCCACACAAATAGTCTGCTACCTCAGCAATATTTTGT 191
    |||||||
Db 121 CATCAGGTGGTAACTCTCCACACAAATAGTCTGCTACCTCAGCAATATTTTGT 180
    |||||||

QY 192 CTGTGTAC 201
    |||||||
Db 181 CTGTGTAC 190
    |||||||

RESULT 5
BI053386/c      346 bp      mRNA      linear      EST 15-JUN-2001
LOCUS        BI053386/c
DEFINITION    RCS-GN0282-250101-011-E07 GN0282 Homo sapiens cDNA, mRNA sequence.
ACCESSION     BI053386
VERSION       BI053386.1 GI:14460916
KEYWORDS      EST.
SOURCE        human.
ORGANISM      Homo sapiens
              Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
              Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE     1 (bases 1 to 346)
              Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
              Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
              Goldman,G.H., Carvalho,A.F., Matsukuma,A., Bala,G.S., Simpson,D.H.,
              Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
              ,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
              Simpson,A.J.
              Shotgun sequencing of the human transcriptome with ORF expressed
              sequence tags
              Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
              2020263
              Contact: Simpson A.J.G.
              Laboratory of Cancer Genetics
              Ludwig Institute for Cancer Research
              Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
              Brazil
              Tel: +55-11-2704922
              Fax: +55-11-2707001
              Email: asimpson@ludwig.org.br
              This sequence was derived from the FAPESP/LICR Human Cancer Genome
              Project. This entry can be seen in the following URL
              (http://www.ludwig.org.br/scripts/gethtml2.pl?tl-RCS5t2-RCS-GN0282-
              250101-011-E07&t3=2001-01-25&t4=1)
              Seq primer: puc 18 forward
              High quality sequence stop: 342.
              Location/Qualifiers
                source
                  1. 346
                    /organism="Homo sapiens"
                    /db_xref="taxon:9606"
                    /clone_id="GN0282"
                    /dev_stage="Adult"
                    /note="Organ: placenta.normal; Vector: puc18; Site_1: SmaI
                    ; Site_2: SmaI; A mini-library was made by cloning
                    products derived from ORESTES PCR (U.S. Letters Patent
                    application No. 196,716 - Ludwig Institute for Cancer
                    Research) profiles into the pUC 18 vector. Reverse
                    transcription of tissue mRNA and cDNA amplification were
                    performed under low stringency conditions."
BASE COUNT      84 a      48 c      116 g      98 t
ORIGIN

Query Match      85.2%; Score 171.2; DB 13; Length 346;
Best Local Similarity 98.3%; Pred. No. 2.9e-43;
Matches 173; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
```

QY 1 AACTTCAGCAGAAATAAACACCACTTCCTTTAGTAGACCTCTGTTCCANTGTG 60
|||||
Db 187 AACTTCAGCAGAAATAAACACCACTTCCTTTAGTAGACCTCTGTTCCANTGTG 128
|||||
QY 61 GAATTAACCCATACCTCAAACTCAGTGTAAATTTAGCAATACATACATACCAACC 120
|||||
Db 127 GAATTAACCCATACCTCAAACTCAGTGTAAATTTAGCAATACATACATACCAACC 68
|||||
QY 121 AACTCCCAATGATCAGTGGGTGAATCTCTCCACACAAATATGCTGCTACCCCTC 176
|||||
Db 67 AACTCCCAATGATCAGTGGGTGAATCTCTCCACACAAATATGCTGCTACCCCTC 12
|||||

RESULT 6
BI055459/c 315 bp mRNA linear EST 15-JUN-2001
LOCUS BI055459
DEFINITION RC5-GN0282-060201-012-H05 GN0282 Homo sapiens cDNA, mRNA sequence.
ACCESSION BI055459
VERSION BI055459.1 GI:14462989
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
AUTHORS Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
Nagal,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
Goldman,G.H., Carvalho,A.F., Matsukuma,A., Bala,G.S., Simpson,D.H.,
Brunstein,A., deoliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
Simpson,A.J.
TITLE Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
MEDLINE 20202663
COMMENT Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?l=RC5&t2=RC5-GN0282-
060201-012-H05&t3=2001-02-06&t4=1)
Seq primer: puc 18 forward
High quality sequence start: 11
High quality sequence stop: 242.
Location/Qualifiers
FEATURES
source
1..315
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="GN0282"
/dev_stage="Adult"
/note="Organ: placenta.normal; Vector: puc18; Site_1: SmaI
; Site_2: SmaI; A mini-library was made by cloning
products derived from ORESTES PCR (U.S. Letters Patent
Application No. 196,716 - Ludwig Institute for Cancer
Research) profiles into the puc 18 vector. Reverse
transcription of tissue mRNA and cDNA amplification were
performed under low stringency conditions."

BASE COUNT 74 a 49 c 95 g 95 t 2 others
ORIGIN
Query Match 85.1%; Score 171; DB 13; Length 315;
Best local Similarity 100.0%; Pred. No. 3.3e-43;
Matches 171; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AACTTCAGCAGAAATAAACACCACTTCCTTTAGTAGACCTCTGTTCCANTGTG 60
|||||
Db 179 AACTTCAGCAGAAATAAACACCACTTCCTTTAGTAGACCTCTGTTCCANTGTG 120
|||||

QY 61 GAATTAACCCATACCTCAAACTCAGTGTAAATTTAGCAATACATACATACCAACC 120
|||||
Db 119 GAATTAACCCATACCTCAAACTCAGTGTAAATTTAGCAATACATACATACCAACC 60
|||||
QY 121 AACTCCCAATGATCAGTGGGTGAATCTCTCCACACAAATATGCTGCTA 171
|||||
Db 59 AACTCCCAATGATCAGTGGGTGAATCTCTCCACACAAATATGCTGCTA 9
|||||

RESULT 7
BI056080/c 315 bp mRNA linear EST 15-JUN-2001
LOCUS BI056080
DEFINITION RC5-GN0282-150201-012-H05 GN0282 Homo sapiens cDNA, mRNA sequence.
ACCESSION BI056080
VERSION BI056080.1 GI:14463610
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
AUTHORS Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
Nagal,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
Goldman,G.H., Carvalho,A.F., Matsukuma,A., Bala,G.S., Simpson,D.H.,
Brunstein,A., deoliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
Simpson,A.J.
TITLE Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
MEDLINE 20202663
COMMENT Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?l=RC5&t2=RC5-GN0282-
150201-012-H05&t3=2001-02-15&t4=1)
Seq primer: puc 18 forward
High quality sequence start: 11
High quality sequence stop: 242.
Location/Qualifiers
FEATURES
source
1..315
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="GN0282"
/dev_stage="Adult"
/note="Organ: placenta.normal; Vector: puc18; Site_1: SmaI
; Site_2: SmaI; A mini-library was made by cloning
products derived from ORESTES PCR (U.S. Letters Patent
Application No. 196,716 - Ludwig Institute for Cancer
Research) profiles into the puc 18 vector. Reverse
transcription of tissue mRNA and cDNA amplification were
performed under low stringency conditions."

BASE COUNT 74 a 49 c 95 g 95 t 2 others
ORIGIN
Query Match 85.1%; Score 171; DB 13; Length 315;
Best local Similarity 100.0%; Pred. No. 3.3e-43;
Matches 171; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AACTTCAGCAGAAATAAACACCACTTCCTTTAGTAGACCTCTGTTCCANTGTG 60
|||||
Db 179 AACTTCAGCAGAAATAAACACCACTTCCTTTAGTAGACCTCTGTTCCANTGTG 120
|||||

QY 61 GAATTAACCCATACCTCAAACTCAGTGTAAATTTAGCAATACATACATACCAACC 120
|||||

QY 179 GAATATTTTGTCTGTG 197
|||||
Db 298 GAATATTTTGTCTGTG 316
|||||

RESULT 10
BI056592/c 345 bp mRNA linear EST 15-JUN-2001
LOCUS RC5-GN0282-210201-013-E03 GN0282 Homo sapiens cDNA, mRNA sequence.
DEFINITION BI056592
ACCESSION BI056592
VERSION BI056592.1 GI:14464122
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
1 (bases 1 to 345)
Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R.,
Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F.,
Goldman, G.H., Carvalho, A.F., Matsukuma, A., Bala, G.S., Simpson, D.H.,
Brunstein, A., de Oliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare,
M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and
Simpson, A.J.
Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)

JOURNAL MEDLINE
COMMENT 20202663
Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?l=RC5&t2=RC5-GN0282-
210201-013-E03&t3=2001-02-21&t4=1)
Seq primer: puc 18 forward
High quality sequence stop: 344.
Location/Qualifiers
1. 345
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="GN0282"
/dev_stage="Adult"
/note="Organ: placenta,normal; Vector: puc18; Site_1: SmaI
; Site_2: SmaI; A mini-library was made by cloning
products derived from ORESTES PCR (U.S. Letters Patent
application No. 196/716 - Ludwig Institute for Cancer
Research) profiles into the puc 18 vector. Reverse
transcription of tissue mRNA and cDNA amplification were
performed under low stringency conditions."

BASE COUNT 85 a 45 c 107 g 108 t
ORIGIN

Query Match 84.3%; Score 169.4; DB 13; Length 345;
Best Local Similarity 99.4%; Pred. No. 1.1e-42;
Matches 170; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AACTTCAGCAGAAATAAACACACTTCGTTTGTAGTAGACCTCTTGTTCATCTG 60
|||||
Db 172 AACTTCAGCAGAAATAAACACACTTCGTTTGTAGTAGACCTCTTGTTCATCTG 113
|||||

QY 61 GAAATTAACCATCTCAACCTCACTCTGTGTAATAATTAGCAATATCTATACACAC 120
|||||
Db 112 GAAATTAACCATCTCAACCTCACTCTGTGTAATAATTAGCAATATCTATACACAC 53
|||||

QY 121 AACTCCATGCATCAGTGGGTAACTCTCCACACAAATAGTGTGCTTA 171
|||||
Db 52 AACTCCATGCATCAGTGGGTAACTCTCCACACAAATAGTGTGCTTA 2
|||||

RESULT 11
BI056596/c 313 bp mRNA linear EST 15-JUN-2001
LOCUS RC5-GN0282-210201-013-H02 GN0282 Homo sapiens cDNA, mRNA sequence.
DEFINITION BI056596
ACCESSION BI056596
VERSION BI056596.1 GI:14464126
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
1 (bases 1 to 313)
Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R.,
Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F.,
Goldman, G.H., Carvalho, A.F., Matsukuma, A., Bala, G.S., Simpson, D.H.,
Brunstein, A., de Oliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare,
M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and
Simpson, A.J.
Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)

JOURNAL MEDLINE
COMMENT 20202663
Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?l=RC5&t2=RC5-GN0282-
210201-013-H02&t3=2001-02-21&t4=1)
Seq primer: puc 18 forward
High quality sequence stop: 312.
Location/Qualifiers
1. 313
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="GN0282"
/dev_stage="Adult"
/note="Organ: placenta,normal; Vector: puc18; Site_1: SmaI
; Site_2: SmaI; A mini-library was made by cloning
products derived from ORESTES PCR (U.S. Letters Patent
application No. 196/716 - Ludwig Institute for Cancer
Research) profiles into the puc 18 vector. Reverse
transcription of tissue mRNA and cDNA amplification were
performed under low stringency conditions."

BASE COUNT 77 a 46 c 99 g 91 t
ORIGIN

Query Match 84.1%; Score 169; DB 13; Length 313;
Best Local Similarity 100.0%; Pred. No. 1.4e-42;
Matches 169; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AACTTCAGCAGAAATAAACACACTTCGTTTGTAGTAGACCTCTTGTTCATCTG 60
|||||
Db 169 AACTTCAGCAGAAATAAACACACTTCGTTTGTAGTAGACCTCTTGTTCATCTG 110
|||||

QY 61 GAAATTAACCATCTCAACCTCACTCTGTGTAATAATTAGCAATATCTATACACAC 120
|||||
Db 109 GAAATTAACCATCTCAACCTCACTCTGTGTAATAATTAGCAATATCTATACACAC 50
|||||

QY 121 AACTCCATGCATCAGTGGGTAACTCTCCACACAAATAGTGTGCTTA 169
|||||
Db 49 AACTCCATGCATCAGTGGGTAACTCTCCACACAAATAGTGTGCTTA 1
|||||

RESULT 12
BI052559/c 373 bp mRNA linear EST 15-JUN-2001
LOCUS

DEFINITION RCS-GN0282-220101-011-A04 GN0282 Homo sapiens cDNA, mRNA sequence.
ACCESSION B1052559
VERSION B1052559.1 GI:14460089
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
AUTHORS Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
Goldman,G.H., Carvalho,A.F., Matsukuma,A., Bala,G.S., Simpson,D.H.,
Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
Simpson,A.J.J.
TITLE Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
MEDLINE 20202663
COMMENT Contact: Simpson A.J.J.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?rl=RCS5<2-RCS-GN0282-
220101-011-A04<3-2001-01-22<4-1)
Seq primer: puc 18 forward
High quality sequence stop: 373.
Location/Qualifiers
1..373
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="GN0282"
/dev_stage="Adult"
/note="Organ: placenta.normal; Vector: puc18; Site_1: Sma1
; Site_2: Sma1; A mini-library was made by cloning
products derived from ORESTES PCR (U.S. Letters Patent
Application No. 196,716 - Ludwig Institute for Cancer
Research) profiles into the puc 18 vector. Reverse
transcription of tissue mRNA and cDNA amplification were
performed under low stringency conditions."
BASE COUNT 89 a 52 c 120 g 112 t
ORIGIN
Query Match 84.1%; Score 169; DB 13; Length 373;
Best Local Similarity 100.0%; Pred. No. 1.5e-42;
Matches 169; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 AACTCAGACAGAAATAAACAACCACTCCGTTTAGTAGAGCACTCTGTTCCAACTCG 60
|||||
DB 169 AACTCAGACAGAAATAAACAACCACTCCGTTTAGTAGAGCACTCTGTTCCAACTCG 110
QY 61 GAAATTAACCAATACCTCAAACTCACTGTGTAAATTTAGCAATACATACACACACC 120
|||||
DB 109 GAAATTAACCAATACCTCAAACTCACTGTGTAAATTTAGCAATACATACACACACC 50
QY 121 AACTCCCAATGATCAGTGGGTAACTCTCCACACAATAATAGTCTGCC 169
|||||
DB 49 AACTCCCAATGATCAGTGGGTAACTCTCCACACAATAATAGTCTGCC 1
RESULT 13 373 bp mRNA linear EST 15-JUN-2001
LOCUS B1052567/c
DEFINITION RCS-GN0282-220101-011-E06 GN0282 Homo sapiens cDNA, mRNA sequence.
ACCESSION B1052567
VERSION B1052567.1 GI:14460097
KEYWORDS EST.

SOURCE human.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
AUTHORS Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
Goldman,G.H., Carvalho,A.F., Matsukuma,A., Bala,G.S., Simpson,D.H.,
Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
Simpson,A.J.J.
TITLE Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
MEDLINE 20202663
COMMENT Contact: Simpson A.J.J.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?rl=RCS5<2-RCS-GN0282-
220101-011-E06<3-2001-01-22<4-1)
Seq primer: puc 18 forward
High quality sequence stop: 373.
Location/Qualifiers
1..373
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="GN0282"
/dev_stage="Adult"
/note="Organ: placenta.normal; Vector: puc18; Site_1: Sma1
; Site_2: Sma1; A mini-library was made by cloning
products derived from ORESTES PCR (U.S. Letters Patent
Application No. 196,716 - Ludwig Institute for Cancer
Research) profiles into the puc 18 vector. Reverse
transcription of tissue mRNA and cDNA amplification were
performed under low stringency conditions."
BASE COUNT 89 a 52 c 120 g 112 t
ORIGIN
Query Match 84.1%; Score 169; DB 13; Length 373;
Best Local Similarity 100.0%; Pred. No. 1.5e-42;
Matches 169; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 AACTCAGACAGAAATAAACAACCACTCCGTTTAGTAGAGCACTCTGTTCCAACTCG 60
|||||
DB 169 AACTCAGACAGAAATAAACAACCACTCCGTTTAGTAGAGCACTCTGTTCCAACTCG 110
QY 61 GAAATTAACCAATACCTCAAACTCACTGTGTAAATTTAGCAATACATACACACACC 120
|||||
DB 109 GAAATTAACCAATACCTCAAACTCACTGTGTAAATTTAGCAATACATACACACACC 50
QY 121 AACTCCCAATGATCAGTGGGTAACTCTCCACACAATAATAGTCTGCC 169
|||||
DB 49 AACTCCCAATGATCAGTGGGTAACTCTCCACACAATAATAGTCTGCC 1
RESULT 14 379 bp mRNA linear EST 15-JUN-2001
LOCUS B1052566/c
DEFINITION RCS-GN0282-220101-011-E05 GN0282 Homo sapiens cDNA, mRNA sequence.
ACCESSION B1052566
VERSION B1052566.1 GI:14460096
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
AUTHORS
1 (bases 1 to 379)
Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M. R., Nagai, M. A., da Silva, W. Jr., Zago, M. A., Bordin, S., Costa, F. F., Goldman, G. H., Carvalho, A. F., Matsukuma, A., Bala, G. S., Simpson, D. H., Brunstein, A., deoliveira, P. S., Bucher, P., Jongeneel, C. V., O'Hare, M. J., Soares, F., Brentani, R. R., Reis, L. F., de Souza, S. J. and Simpson, A. J.

TITLE
Shotgun sequencing of the human transcriptome with ORF expressed sequence tags

JOURNAL
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)

COMMENT
Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=RC5&ct=RC5-GN0282-220101-011-E05&t3=2001-01-22&t4=1)
Seq primer: puc 18 forward
High quality sequence stop: 363.
Location/Qualifiers
1..379
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_id="GN0282"
/dev_stage="Adult"
/note="Organ: placenta.normal; Vector: puc18; Site_1: SmaI
; Site_2: SmaI; A mini-library was made by cloning
products derived from ORESTES PCR (U.S. Letters Patent
Application No. 196,716 - Ludwig Institute for Cancer
Research) profiles into the puc 18 vector. Reverse
transcription of tissue mRNA and cDNA amplification were
performed under low stringency conditions."

BASE COUNT
ORIGIN
90 a 53 c 123 g 112 t 1 others

Query Match 84.1%; Score 169; DB 13; Length 379;
Best Local Similarity 100.0%; Pred. No. 1.5e-42;
Matches 169; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AACTCAGCAGCAAGAAATTAACACACCTTCCTTTAGTAGAGCTCTGTGTTCCAAATCG 60
|||||
Db 169 AACTCAGCAGCAAGAAATTAACACACCTTCCTTTAGTAGAGCTCTGTGTTCCAAATCG 110

QY 61 GAATAACCATACCTCAAACTCAGCTGCTGTAAATTTAGCAATACTACTACATACACAACC 120
|||||
Db 109 GAATAACCATACCTCAAACTCAGCTGCTGTAAATTTAGCAATACTACTACATACACAACC 50

QY 121 AACTCCCAATGATCAGGTGGGTAACTCTCCACACAAATAGTCTGCC 169
|||||
Db 49 AACTCCCAATGATCAGGTGGGTAACTCTCCACACAAATAGTCTGCC 1

RESULT 15
BIOS2569/c 382 bp mRNA linear EST 15-JUN-2001
LOCUS RC5-GN0282-220101-011-G04 GN0282 Homo sapiens cDNA, mRNA sequence.
DEFINITION
ACCESSION BI052569
VERSION BI052569.1 GI:14460099
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE
1 (bases 1 to 382)
Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M. R., Nagai, M. A., da Silva, W. Jr., Zago, M. A., Bordin, S., Costa, F. F., Goldman, G. H., Carvalho, A. F., Matsukuma, A., Bala, G. S., Simpson, D. H.,

Brunstein, A., deoliveira, P. S., Bucher, P., Jongeneel, C. V., O'Hare, M. J., Soares, F., Brentani, R. R., Reis, L. F., de Souza, S. J. and Simpson, A. J.

TITLE
Shotgun sequencing of the human transcriptome with ORF expressed sequence tags

JOURNAL
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)

COMMENT
Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=RC5&ct=RC5-GN0282-220101-011-G04&t3=2001-01-22&t4=1)
Seq primer: puc 18 forward
High quality sequence stop: 381.
Location/Qualifiers
1..382
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_id="GN0282"
/dev_stage="Adult"
/note="Organ: placenta.normal; Vector: puc18; Site_1: SmaI
; Site_2: SmaI; A mini-library was made by cloning
products derived from ORESTES PCR (U.S. Letters Patent
Application No. 196,716 - Ludwig Institute for Cancer
Research) profiles into the puc 18 vector. Reverse
transcription of tissue mRNA and cDNA amplification were
performed under low stringency conditions."

BASE COUNT
ORIGIN
92 a 55 c 121 g 114 t

Query Match 84.1%; Score 169; DB 13; Length 382;
Best Local Similarity 100.0%; Pred. No. 1.5e-42;
Matches 169; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AACTCAGCAGCAAGAAATTAACACACCTTCCTTTAGTAGAGCTCTGTGTTCCAAATCG 60
|||||
Db 169 AACTCAGCAGCAAGAAATTAACACACCTTCCTTTAGTAGAGCTCTGTGTTCCAAATCG 110

QY 61 GAATAACCATACCTCAAACTCAGCTGCTGTAAATTTAGCAATACTACTACATACACAACC 120
|||||
Db 109 GAATAACCATACCTCAAACTCAGCTGCTGTAAATTTAGCAATACTACTACATACACAACC 50

QY 121 AACTCCCAATGATCAGGTGGGTAACTCTCCACACAAATAGTCTGCC 169
|||||
Db 49 AACTCCCAATGATCAGGTGGGTAACTCTCCACACAAATAGTCTGCC 1

Search completed: May 2, 2003, 16:21:32
Job time: 1115 secs

GenCore version 5.1.5
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 2, 2003, 14:52:21 ; Search time 39.333 Seconds

(without alignments)
1567.170 Million cell updates/sec

Title: US-09-719-554-3_COPY_8500_8700

Perfect score: 201
Sequence: 1 aactcagccagcaataaa.....tattttgtctgtgtracc 201

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

Issued_patents_NA:*
1: /cgn2_6/p10data/2/1na/5A.COMB.seq:*
2: /cgn2_6/p10data/2/1na/5B.COMB.seq:*
3: /cgn2_6/p10data/2/1na/6A.COMB.seq:*
4: /cgn2_6/p10data/2/1na/6B.COMB.seq:*
5: /cgn2_6/p10data/2/1na/PCTUS.COMB.seq:*
6: /cgn2_6/p10data/2/1na/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	201	100.0	2946	US-09-175-928-3	Sequence 3, Appli
2	29	14.4	1728	US-08-229-781-28	Sequence 28, Appl
3	29	14.4	1728	US-08-630-918-28	Sequence 28, Appl
4	29	14.4	1728	US-09-004-422-28	Sequence 28, Appl
5	28.2	14.0	465	PCT-US96-05320A-10	Sequence 10, Appl
6	27.8	13.8	1010	US-09-453-702B-4	Sequence 4, Appli
7	27.8	13.8	1857	US-09-134-001C-332	Sequence 332, App
8	27.8	13.8	2248	US-08-639-237-1	Sequence 1, Appli
9	27.8	13.8	2248	US-08-798-405-1	Sequence 1, Appli
10	27.8	13.8	2264	US-09-167-109-6	Sequence 6, Appli
11	27.8	13.8	4343	US-08-377-690-2	Sequence 2, Appli
12	27.8	13.8	4360	US-08-486-380-1	Sequence 1, Appli
13	27.8	13.8	4360	US-08-436-080-1	Sequence 1, Appli
14	27.8	13.8	4360	US-08-250-848-1	Sequence 1, Appli
15	27.8	13.8	4360	US-08-438-192-1	Sequence 1, Appli
16	27.8	13.8	4360	US-08-475-971-1	Sequence 1, Appli
17	27.8	13.8	4360	US-09-021-203-1	Sequence 1, Appli
18	27.8	13.8	4767	US-09-532-584B-1	Sequence 1, Appli
19	27.8	13.8	8697	US-08-961-527-123	Sequence 123, App
20	27.8	13.8	65042	US-07-784-316-3	Sequence 3, Appli
21	27.8	13.8	99500	US-09-798-096-10	Sequence 10, Appli
22	27.6	13.7	422	US-08-470-179-73	Sequence 73, Appli
23	27.4	13.6	1135	US-08-229-781-49	Sequence 49, Appli
24	27.4	13.6	1135	US-08-630-918-49	Sequence 49, Appli
25	27.4	13.6	1135	US-09-004-422-49	Sequence 49, Appli
26	27.4	13.6	1135	US-08-229-781-46	Sequence 46, Appli
27	27.4	13.6	1783	US-08-630-918-46	Sequence 46, Appli

28	27.4	13.6	1783	US-09-004-422-46	Sequence 46, Appli
29	27.4	13.6	1947	US-08-604-991-1	Sequence 1, Appli
30	27.4	13.6	1947	US-09-363-639-1	Sequence 1, Appli
31	27	13.4	1569	US-09-134-001C-918	Sequence 918, App
32	27	13.4	7244	US-08-378-313-26	Sequence 26, Appli
33	26.8	13.3	394	US-08-229-781-36	Sequence 36, Appli
34	26.8	13.3	394	US-08-630-918-36	Sequence 36, Appli
35	26.8	13.3	730	US-09-004-422-36	Sequence 36, Appli
36	26.8	13.3	730	US-07-826-943A-10	Sequence 10, Appli
37	26.8	13.3	1000	US-07-960-112B-1	Sequence 1, Appli
38	26.8	13.3	1000	US-07-960-112B-3	Sequence 3, Appli
39	26.8	13.3	1000	US-08-301-316B-1	Sequence 1, Appli
40	26.8	13.3	1000	US-08-301-316B-3	Sequence 3, Appli
41	26.8	13.3	1000	US-08-473-389B-1	Sequence 1, Appli
42	26.8	13.3	1000	US-08-473-389B-3	Sequence 3, Appli
43	26.8	13.3	1000	US-08-853-831-1	Sequence 1, Appli
44	26.8	13.3	1000	US-08-853-831-3	Sequence 3, Appli
45	26.8	13.3	1000	PCT-US93-09774-1	Sequence 1, Appli

ALIGNMENTS

```
RESULT 1
US-09-175-928-3
: Sequence 3, Application US/09175928A
: Patent No. 6312921
: GENERAL INFORMATION:
: APPLICANT: Jacobs, Kenneth
: APPLICANT: McCoy, John M.
: APPLICANT: Lavallie, Edward R.
: APPLICANT: Collins-Racie, Lisa A.
: APPLICANT: Evans, Cheryl
: APPLICANT: Werberg, David
: APPLICANT: Treacy, Maurice
: APPLICANT: Mi, Sha
: APPLICANT: Genetics Institute, Inc.
: TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES ENCODING THEM
: FILE REFERENCE: 6006B, A172A
: CURRENT APPLICATION NUMBER: US/09/175, 928A
: CURRENT FILING DATE: 1998-10-20
: NUMBER OF SEQ ID NOS: 62
: SOFTWARE: PatentIn Ver. 2.0
: SEQ ID NO 3
: LENGTH: 2946
: TYPE: DNA
: ORGANISM: Homo sapiens
:
Query Match      100.0%: Score 201; DB 4; Length 2946;
Best Local Similarity 100.0%: Pred. No. 2.8e-52;
Matches 201; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
:
OY 1 AACTTCAGCACGAAATTAACACCACTCCGTTTGTAGTAGACCTCTGTTCCATCTG 60
    |||||||
DB 1549 AACTTCAGCACGAAATTAACACCACTCCGTTTGTAGTAGACCTCTGTTCCATCTG 1608
    |||||||
OY 61 GAAATTAACCCATCTCAACCTCACTGCTGTAATAATTGACATACATACACAC 120
    |||||||
DB 1609 GAAATTAACCCATCTCAACCTCACTGCTGTAATAATTGACATACATACACAC 1668
    |||||||
OY 121 AACTCCCAATGATAGGTGGGTACTCTCCCAACAAATAGTCTGCTTACCCAGGA 180
    |||||||
DB 1669 AACTCCCAATGATAGGTGGGTACTCTCCCAACAAATAGTCTGCTTACCCAGGA 1728
    |||||||
OY 181 AATTTTGTCTGTGTGTAC 201
    |||||||
DB 1729 AATTTTGTCTGTGTGTAC 1749
    |||||||
RESULT 2
US-08-229-781-28
: Sequence 28, Application US/08229781
```

Patent No. 5589174
GENERAL INFORMATION:
APPLICANT: Yoshinobu OKUNO et al.
TITLE OF INVENTION: ANTI-HUMAN INFLUENZA VIRUS ANTIBODY
NUMBER OF SEQUENCES: 58
CORRESPONDENCE ADDRESS:
ADDRESSEE: Wenderoth, Lind & Ponack
STREET: 805 Fifteenth Street, N.W., #700
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20005
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.5 inch, 1.44 mb
COMPUTER: IBM Compatible
OPERATING SYSTEM: MS-DOS
SOFTWARE: Wordperfect 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/229,781
FILING DATE: April 19, 1994
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/054,016
FILING DATE: April 29, 1993
ATTORNEY/AGENT INFORMATION:
NAME: Warren M. Cheek, Jr.
REGISTRATION NUMBER: 33,367
REFERENCE/DOCKET NUMBER:
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-371-8850
TELEFAX:
TELEX:
INFORMATION FOR SEQ ID NO: 28:
SEQUENCE CHARACTERISTICS:
LENGTH: 1728 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA to genomic RNA
HYPOTHETICAL:
ANTI-SENSE:
FRAGMENT TYPE:
ORIGINAL SOURCE:
ORGANISM: A/12um1/5/65
STRAIN:
INDIVIDUAL ISOLATE:
DEVELOPMENTAL STAGE:
HAPLOTYPE:
TISSUE TYPE:
CELL TYPE:
CELL LINE:
ORGANELLE:
IMMEDIATE SOURCE:
LIBRARY:
CLONE:
POSITION IN GENOME:
CHROMOSOME/SEGMENT:
MAP POSITION:
UNITS:
FEATURE:
NAME/KEY:
LOCATION:
IDENTIFICATION METHOD:
OTHER INFORMATION:
PUBLICATION INFORMATION:
AUTHORS:
TITLE:
JOURNAL:
VOLUME:
ISSUE:
PAGES:
DATE:
DOCUMENT NUMBER:

FILING DATE:
PUBLICATION DATE:
RELEVANT RESIDUES IN SEQ ID NO:
US-08-229-781-28
Query Match 14.4%; Score 29; DB 1; Length 1728;
Best Local Similarity 50.4%; Pred. No. 6;
Matches 71; Conservative 0; Mismatches 70; Indels 0; Gaps 0;
QY 3 CTTCAGCAGAAATAACACCACTCCGTTTAGTAGACACTTGTTCATCTGGA 62
DB 906 CTTTGGAGCAATPAATACACACTCTTTTCACATGTCACACCACTACATAGTG 965
QY 63 AATAACCCATACCTCAACCTCACCTGTAAATTAGCAATACATACACCA 122
DB 966 AATGCCCAATATGTAAATCGAGAAATGTGCTTAGCAGACAGACTAAGAAATGTC 1025
QY 123 CTCCCAATGCATCAGTGGT 143
DB 1026 CCCAGATTGAATCAAGAGAT 1046
RESULT 3
US-08-630-918-28
Sequence 28, Application US/08630918
Patent No. 561350
GENERAL INFORMATION:
APPLICANT: Yoshinobu OKUNO et al.
TITLE OF INVENTION: ANTI-HUMAN INFLUENZA VIRUS ANTIBODY
NUMBER OF SEQUENCES: 58
CORRESPONDENCE ADDRESS:
ADDRESSEE: Wenderoth, Lind & Ponack
STREET: 805 Fifteenth Street, N.W., #700
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20005
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.5 inch, 1.44 mb
COMPUTER: IBM Compatible
OPERATING SYSTEM: MS-DOS
SOFTWARE: Wordperfect 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/630,918
FILING DATE: April 5, 1996
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/229,781
FILING DATE: April 19, 1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/054,016
FILING DATE: April 29, 1993
ATTORNEY/AGENT INFORMATION:
NAME: Warren M. Cheek, Jr.
REGISTRATION NUMBER: 33,367
REFERENCE/DOCKET NUMBER:
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-371-8850
TELEFAX:
TELEX:
INFORMATION FOR SEQ ID NO: 28:
SEQUENCE CHARACTERISTICS:
LENGTH: 1728 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA to genomic RNA
ORIGINAL SOURCE:
ORGANISM: A/12um1/5/65
US-08-630-918-28
Query Match 14.4%; Score 29; DB 1; Length 1728;
Best Local Similarity 50.4%; Pred. No. 6;

```

Matches 71; Conservative 0; Mismatches 70; Indels 0; Gaps 0
QY 3 CTTCAGCACAATAAACAACACCTCCGTTTATAGAGACTCTTGTTCATCTGA 62
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 906 CTTTGGGGAGCAATAATCAACACTACTCTTTTCACAAATGTCACCCACACTGACAATAGGTG 965
QY 63 AATAACCATATCCTCAACCTCACCCTGTGTAAATTTTGCATTACTATACACAACCA 122
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 966 AATGCCCAATATGTATAAATCGGGAATTTGTCTTAGCACAAGAGACTAAGGAATGTTTC 1025
QY 123 CTCCCAATGCATCAGTGGGT 143
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1026 CCCAGATTGAATCAAGAGGAT 1046

RESULT 4
US-09-004-422-28
; Sequence 28, Application US/09004422
; Patent No. 6337070
; GENERAL INFORMATION:
; APPLICANT: Yoshinobu OKUNO et al.
; TITLE OF INVENTION: POLYPEPTIDES FOR USE IN GENERATING
; TITLE OF INVENTION: ANTI-HUMAN INFLUENZA VIRUS ANTIBODIES (AS AMENDED)
; NUMBER OF SEQUENCES: 58
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Menderoth, Lind & Ponack, L.L.P.
; STREET: 2033 K Street, N.W., #800
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20006
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 inch, 1.44 mb
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: Wordperfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/004,422
; FILING DATE: January 8, 1998
; CLASSIFICATION: 530
; PRIOR APPLICATION NUMBER:
; APPLICATION NUMBER: 08/443,862
; FILING DATE: May 22, 1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/229,781
; FILING DATE: April 19, 1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/054,016
; FILING DATE: April 29, 1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Warren M. Cheek, Jr.
; REGISTRATION NUMBER: 33,367
; REFERENCE/DOCKET NUMBER:
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-721-8200
; TELEFAX: 202-721-8250
; TELEX:
; INFORMATION FOR SEQ ID NO: 28:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1728 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA to genomic RNA
; HYPOTHEetical:
; ANTI-SENSE:
; FRAGMENT TYPE:
; ORIGINAL SOURCE:
; ORGANISM: A/12um1/5/65
; STRAIN:
; INDIVIDUAL ISOLATE:
; DEVELOPMENTAL STAGE:
; HAPLOTYPE:

```

```

TISSUE TYPE:
CELL TYPE:
CELL LINE:
ORGANELLE:
IMMEDIATE SOURCE:
LIBRARY:
CLONE:
POSITION IN GENOME:
CHROMOSOME/SEGMENT:
MAP POSITION:
UNITS:
FEATURE:
NAME/KEY:
LOCATION:
IDENTIFICATION METHOD:
OTHER INFORMATION:
PUBLICATION INFORMATION:
AUTHORS:
TITLE:
JOURNAL:
VOLUME:
ISSUE:
PAGES:
DATE:
DOCUMENT NUMBER:
FILING DATE:
PUBLICATION DATE:
RELEVANT RESIDUES IN SEQ ID NO:
US-09-004-422-28

Query Match          14.4%; Score 29; DB 4; Length 1728;
Best Local Similarity 50.4%; Pred. No. 6;
Matches 71; Conservative 0; Mismatches 70; Indels 0; Gaps 0;

QY 3 CTTGACACAGAAATAAACACCACCTCCGTTTGTAGAGACCCTGTGTTCCAACTGGA 62
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 906 CTTGGGGAGCAATAATACACACTACCTTTTCACATGTCCACCACGTGACATAGTG 965
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 63 AATAACCATACCTCAACCTGCCTGTTAAATTTAGCAATACTACATACACAA 122
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 966 AATGCCCAATATGTAAATGCGAATAATGTGCTTACCAACAGACTAAGAAATGTT 1025
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 123 CTCCCAATGATCAGGTGGT 143
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1026 CCCAGATTGAATCAAGAGGAT 1046
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 5
PCT-US96-05320A-10/c
; Sequence 10, Application PC/TUS9605320A
; GENERAL INFORMATION:
; APPLICANT: Human Genome Sciences
; APPLICANT: 9410 Key West Avenue
; APPLICANT: Rockville, MD 20850
; APPLICANT: United States of America
; APPLICANT: Johns Hopkins University
; APPLICANT: 720 Rutland Avenue
; APPLICANT: Baltimore, MD 21205
; APPLICANT: United States of America
; APPLICANT: Mark D. Adams
; APPLICANT: Owen White
; APPLICANT: Hamilton O. Smith
; APPLICANT: J. Craig Venter
; TITLE OF INVENTION: Nucleotide Sequence of the Haemophilus Influenzae Rd Genom
; NUMBER OF SEQUENCES: 48
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sterne, Kessler, Goldstein & Fox
; STREET: 1100 New York Avenue, Suite 600
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20003-3934
; COMPUTER READABLE FORM:

```

```

; MEDIUM TYPE: Diskette, 3.50 inch, 1.44mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US96/05320A
; FILING DATE: April22, 1996
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/476,102
; FILING DATE: June 7, 1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/487,429
; FILING DATE: June 7, 1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Eric K. Steffe
; REGISTRATION NUMBER: 36,688
; REFERENCE/DOCKET NUMBER: 1488,014PC01
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 371-2600
; TELEFAX: (202) 371-2540
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 465 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; PCT-US96-05320A-10

```

	Query Match	Similarity	Score	DB	Length
14	Best Local	53.1%	28.2	5	465
	Matches	60	Pred. No.	7.4	
			Mismatches	53	
			Indels	0	
			Gaps	0	
QY	20	ACACCACTCGCTTTAGTAGACCTCTTGTTCCAACTGGAATAACCCATACCTCA	79		
Db	317	ACACCTTTTCCCTTAATTAAGATTAATTTCTTAACAAATTAATTTGGCAATATCCACA	258		
QY	80	ACCTCACTGTGTAAATTTAGCAATACTACATACACAACCACTCCCAATGC	132		
Db	257	CCTGATATAGTAGGCAAAATCGCATATATAAACAAGTCGCTTCATCCCAATGC	205		

RESULT 6
US-09-453-702B-4/C
Sequence 4, Application US/09453702B
Patent No. 6365723
GENERAL INFORMATION:
APPLICANT: Blatner, Frederick R.
Burland, Valerie
Perna, Nicole T.
Plunkett, Guy
Welch, Rod
TITLE OF INVENTION: No. 6365723el Sequences of E. coli O157
NUMBER OF SEQUENCES: 265
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Quarles & Brady
STREET: 1 South Pinckney Street
CITY: Madison
STATE: WI
COUNTRY: US
ZIP: 53701-2113
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 Inch. 1.44mb storage
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Word Perfect 8.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/453,702B
FILING DATE: 03-Dec-1999
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/110,955
FILING DATE: 04-DEC-1998

```

: ATTORNEY/AGENT INFORMATION:
: NAME: Seay, Nicholas J.
: REGISTRATION NUMBER: 27386
: REFERENCE/DOCKET NUMBER: 960296.95017
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (608) 251-5000
: TELEFAX: (608) 251-9166
: INFORMATION FOR SEQ ID NO: 4:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 1010
: TYPE: nucleic acid
: STRANDEDNESS: double
: TOPOLOGY: linear
: MOLECULE TYPE: DNA (genomic)
: SEQUENCE DESCRIPTION: SEQ ID NO: 4:
US-09-453-702B-4
```

[illegible]

```

RESULT 7
US-09-134-001C-332/C
: Sequence 332, Application US/09134001C
: Patent No. 6380370
: GENERAL INFORMATION:
: APPLICANT: Lynn Doucette-Stamm et al
: TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
: TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
: FILE REFERENCE: GFC-007
: CURRENT APPLICATION NUMBER: US/09/134,001C
: PRIOR FILING DATE: 1998-08-13
: PRIOR APPLICATION NUMBER: US 60/064,964
: PRIOR FILING DATE: 1997-11-08
: PRIOR APPLICATION NUMBER: US 60/055,779
: PRIOR FILING DATE: 1997-08-14
: NUMBER OF SEQ ID NOS: 5674
: SEQ ID NO 332
: LENGTH: 1857
: TYPE: DNA
: ORGANISM: Staphylococcus epidermidis
US-09-134-001C-332

```

```

Query Match Similarity      13.8%; Score 27.8; DB 4; Length 1857;
Best Local Similarity      53.2%; Pred. No. 14;
Matches 59; Conservative 0; Mismatches 52; Indels 0; Gaps 0.

QY      27  TTCGGTTTAGTAGACCTCTGTGTTCCAACTCGAATAACCATPACCTCAACCTCAG 86
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      919  TTCCCTATGAGTAGTATCGTTTAAAGCCACCTAGTGATCTACTCAACTCGAAGTCCAC 860

QY      87  CTGTGTAATAATTGACATPACTACATPACACCAACGACCCCAATGCATCAG 137
      ||| | ||||| | ||||| ||||| ||||| ||||| ||||| |||||
Db      859  CTGAGGTTAATTCTAACTACCATCTTCAGGTAATAAATCTCAATTGATCTG 809

RESULT 8
US-08-639-237-1/c
Sequence 1, Application US/08639237
Patent No. 5710013
GENERAL INFORMATION:
APPLICANT: Goeddel, David V.
APPLICANT: Xiong, Jesse
TITLE OF INVENTION: No. 5710013el protein - TRAF6

```



```

: CLASSIFICATION: 435
:
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 08/639,237
:
: FILING DATE:
:
: ATTORNEY/AGENT INFORMATION:
: NAME: Brezner, David J
:
: REGISTRATION NUMBER: 24,7747
: REFERENCE/DOCKET NUMBER: 796-004/463312
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (415) 781-1989
: TELEFAX: (415) 398-3249
: INFORMATION FOR SEQ ID NO: 1:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 2248 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: CDNA
: FEATURE:
: NAME/KEY: CDS
: LOCATION: *230..1795
:
: US-08-975-405-1
:
: Query Match 13.8% Score 27.8; DB 1; Length 2248;
: Best Local Similarity 55.8%; Pred No. 15;
: Matches 53; Conservative 0; Mismatches 42; Indels 0; Gaps 0

```

Query Match	13.8%	Score 27.8	DB 1	Length 2248
Best Local Similarity	55.8%	Pred. No. 15		
Matches	53	Conservative	0	Mismatches 42; Indels 0; Gaps 0
Oy	103	AATACTACATACACACCAACTCCCAATGATGCATGAGTGGTGAATCCTCCACACAAATA	162	
Db	1766	AACCTTCCTCCGGAAGGTCAACCCATGTCACAAACGGGTGGAGACCTCACAGCCACATA	1707	
Oy	163	GTCTGCTACCCCTCAGAAATATTTTGTGCTGTGG	197	
Db	1706	ATGTGTATCCTTAATGAAAGTTCCTTTGTCTTAGG	1672	

```

RESULT 10
US-09-167-109-6/c
; Sequence 6, Application US/09167109
; Patent No. 6399297
; GENERAL INFORMATION:
; APPLICANT: Baker, Brenda F.
; APPLICANT: Cowser, Lex M.
; APPLICANT: Monia, Brett P.
; APPLICANT: Xu, Xiaoxing S.
; TITLE OF INVENTION: ANTISENSE MODULATION OF TRAF EXPRESSION
; FILE REFERENCE: ISPH-0321
; CURRENT APPLICATION NUMBER: US/09/167,109
; CURRENT FILING DATE: 1998-10-06
; NUMBER OF SEQ ID NOS: 228
; SEQ ID NO 6
; LENGTH: 2264
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (222)..(1790)
; PUBLICATION INFORMATION:
; DATABASE ACCESSION NUMBER: U78798
; DATABASE ENTRY DATE: 1996-12-12
; US-09-167-109-6

```


RESULT 13

US-08-436-080-1
; Sequence 1, Application US/08436080
; Patent No. 5834292
; GENERAL INFORMATION:
; APPLICANT: Rangan, Thirumale S.
; TITLE OF INVENTION: Method for Producing Somaclonal Variant
; TITLE OF INVENTION: Cotton Plants
; NUMBER OF SEQUENCES: 1
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Christie, Parker & Hale
; STREET: P.O. Box 7068
; CITY: Pasadena
; STATE: CA
; COUNTRY: USA
; ZIP: 91109-7068
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentln Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/436,080
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Sharp Esq., Janice A.
; REGISTRATION NUMBER: 34,051
; REFERENCE/DOCKET NUMBER: P114:25705
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (818)795-5843
; TELEFAX: (818)577-1769
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4360 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHEICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Bacillus thuringiensis
US-08-436-080-1

Query Match 13.8%; Score 27.8; DB 2; Length 4360;

Best Local Similarity 49.0%; Pred. No. 18; Matches 74; Conservative 0; Mismatches 77; Indels 0; Gaps 0;

QY 18 AAACACCACTTCGTTTGTAGACGCTCTGTTTCAATCTGGAATAACCATACCTC 77
DB 4169 AAAAAGCGGTGAATTTTGAATATATTTTTCGATTTGGAAGTAACCTTGTGTA 4228
QY 78 AAACCTACCTGTGTAATAATTAGCAATACATACACAAACCACTCCCATGATCAG 137
DB 4229 AAACATCAGCATTTCAAGTCAGACATCAGATATTTTCAAGATCCGTAATTTAGATG 4288
QY 138 GTGGTAACCTCTCCACACAAATAGTCTGC 168
DB 4289 CGACATTTTCCAAAGTACCGAAACATTTAGC 4319

RESULT 14

US-08-250-848-1

; Sequence 1, Application US/08250848
; Patent No. 5856177
; GENERAL INFORMATION:
; APPLICANT: Hudspeth, Richard L.
; TITLE OF INVENTION: PROMOTERS DERIVED FROM THE MAIZE
; TITLE OF INVENTION: PHOSPHOENOLPYRUVATE CARBOXYLASE GENE INVOLVED IN C4
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Christie, Parker & Hale
; STREET: P.O. Box 7068
; CITY: Pasadena
; STATE: CA
; COUNTRY: USA
; ZIP: 91109-7068
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentln Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/250,848
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Sharp Esq., Janice A.
; REGISTRATION NUMBER: 34,051
; REFERENCE/DOCKET NUMBER: P114:25992/JAS
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (818) 577-1769
; TELEFAX: (818) 795-5843
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4360 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHEICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Bacillus thuringiensis
US-08-250-848-1

Query Match 13.8%; Score 27.8; DB 2; Length 4360;

Best Local Similarity 49.0%; Pred. No. 18; Matches 74; Conservative 0; Mismatches 77; Indels 0; Gaps 0;

QY 18 AAACACCACTTCGTTTGTAGACGCTCTGTTTCAATCTGGAATAACCATACCTC 77
DB 4169 AAAAAGCGGTGAATTTTGAATATATTTTTCGATTTGGAAGTAACCTTGTGTA 4228
QY 78 AAACCTACCTGTGTAATAATTAGCAATACATACACAAACCACTCCCATGATCAG 137
DB 4229 AAACATCAGCATTTCAAGTCAGACATCAGATATTTTCAAGATCCGTAATTTAGATG 4288
QY 138 GTGGTAACCTCTCCACACAAATAGTCTGC 168
DB 4289 CGACATTTTCCAAAGTACCGAAACATTTAGC 4319

RESULT 15

US-08-438-192-1
; Sequence 1, Application US/08438192
; Patent No. 5859321
; GENERAL INFORMATION:
; APPLICANT: Rangan, Thirumale S.
; APPLICANT: Anderson, David M.
; APPLICANT: Rajasekaran, Kanniah
; TITLE OF INVENTION: Cotton Somaclonal Variants
; NUMBER OF SEQUENCES: 1

CORRESPONDENCE ADDRESS:
ADDRESSEE: Christie, Parker & Hale
STREET: P. O. Box 7068
CITY: Pasadena
STATE: CA
COUNTRY: USA
ZIP: 91109-7068
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/438,192
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/122,353
FILING DATE:
APPLICATION NUMBER: US 07/122,200
FILING DATE: 18-NOV-1987
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/680,048
FILING DATE: 29-MAR-1991
ATTORNEY/AGENT INFORMATION:
NAME: Sharp Esq., Janice A.
REGISTRATION NUMBER: 34,051
REFERENCE/DOCKET NUMBER: P114:25703
TELECOMMUNICATION INFORMATION:
TELEPHONE: (818)795-5843
TELEFAX: (818)577-1769
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 4360 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Bacillus thuringiensis

GenCore version 5.1.5
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 2, 2003, 16:21:41 ; Search time 63.3333 Seconds

(without alignments)
3750.092 Million cell updates/sec

Title: US-09-719-554-3_COPY_8500_8700

Perfect score: 201

Sequence: 1 aacttcagcagcagaataa.....tatttttctgtgttacc 201

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 746064 seqs, 590810554 residues

Total number of hits satisfying chosen parameters: 1492128

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Maximum Match 0%

Listing first 45 summaries

Database : Published.Applications_NA.*

1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
10: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
11: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
12: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*
13: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
14: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	201	100.0	2930	10 US-09-902-535-1	Sequence 1, Appl
2	201	100.0	2946	9 US-10-114-893-134	Sequence 134, App
3	193	96.0	1481	7 US-08-979-847-105	Sequence 105, App
4	175.4	87.3	792	10 US-09-864-761-21192	Sequence 21192, A
5	175.4	87.3	1894	10 US-09-864-761-4444	Sequence 4444, A
6	33.4	16.6	3457	10 US-09-880-107-2185	Sequence 2185, Ap
7	32.4	16.1	1084	9 US-09-938-842A-3786	Sequence 3786, Ap
8	32.4	16.1	2000	10 US-09-887-576-143	Sequence 143, App
9	30.6	15.2	3309400	9 US-09-738-626-1	Sequence 1, Appl
10	30	14.9	30	10 US-09-902-535-5	Sequence 5, Appl
11	30	14.9	580	10 US-09-822-830A-567	Sequence 567, App
12	30	14.9	1962	10 US-09-822-830A-123	Sequence 123, App
13	29.4	14.6	1024	10 US-09-925-301-15	Sequence 15, Appl
14	29.4	14.6	1999	7 US-08-781-986A-289	Sequence 289, App
15	29.4	14.6	684973	10 US-09-263-959-1	Sequence 1, Appl
16	29.2	14.5	6259	9 US-09-764-891-9689	Sequence 9689, Ap
17	29.2	14.5	6260	9 US-09-764-891-9690	Sequence 9690, Ap
18	29	14.4	1728	10 US-09-918-568-28	Sequence 28, Appl
19	29	14.4	73308	10 US-09-954-456-2276	Sequence 2276, App

C 20	29	14.4	173808	12 US-10-003-806-10	Sequence 10, Appl
C 21	28.6	14.2	659158	9 US-09-771-208-20	Sequence 20, Appl
C 22	28.4	14.1	369	10 US-09-983-965-110	Sequence 110, App
C 23	28.2	14.0	1136	10 US-09-849-967A-4	Sequence 4, Appl
C 24	28.2	14.0	1689	10 US-09-849-967A-1	Sequence 1, Appl
C 25	28	13.9	805	9 US-09-775-049-27	Sequence 27, Appl
C 26	28	13.9	958	10 US-09-864-761-9654	Sequence 9654, Ap
C 27	28	13.9	1368	9 US-09-910-186A-17	Sequence 17, Appl
C 28	27.8	13.8	1010	9 US-10-114-170-4	Sequence 4, Appl
C 29	27.8	13.8	2264	9 US-10-067-125-6	Sequence 6, Appl
C 30	27.8	13.8	4360	10 US-09-756-643-1	Sequence 1, Appl
C 31	27.8	13.8	7214	9 US-10-205-942-1	Sequence 1, Appl
C 32	27.8	13.8	34641	10 US-09-954-456-11110	Sequence 1110, Ap
C 33	27.8	13.8	34641	10 US-09-954-456-1187	Sequence 1787, Ap
C 34	27.8	13.8	65042	9 US-10-229-124-3	Sequence 3, Appl
C 35	27.6	13.7	247	10 US-09-783-590-10685	Sequence 10685, A
C 36	27.6	13.7	500	10 US-09-783-590-5616	Sequence 5616, Ap
C 37	27.4	13.6	389	9 US-09-918-568-95	Sequence 95, Appl
C 38	27.4	13.6	462	10 US-09-560-863-95	Sequence 6, Appl
C 39	27.4	13.6	801	10 US-09-921-823-6	Sequence 49, Appl
C 40	27.4	13.6	1135	10 US-09-918-568-49	Sequence 46, Appl
C 41	27.4	13.6	1783	10 US-09-918-568-46	Sequence 156, App
C 42	27.2	13.5	342	9 US-09-933-797-156	Sequence 1227, Ap
C 43	27.2	13.5	2867	10 US-09-954-456-1327	Sequence 2354, Ap
C 44	27.2	13.5	2867	10 US-09-880-107-2354	Sequence 327, App
C 45	27.2	13.5	6005	10 US-09-070-927A-327	

ALIGNMENTS

RESULT 1
US-09-902-535-1
; Sequence 1, Application US/09902535
; Patent No. US20020102530A1
; GENERAL INFORMATION:
; APPLICANT: Keith, Jr., James C.
; APPLICANT: McCoy, John M.
; APPLICANT: MI, Sina
; TITLE OF INVENTION: Methods and compositions for diagnosing
; TITLE OF INVENTION: and treating preeclampsia and gestational trophoblast
; FILE REFERENCE: GIN-6006B4
; CURRENT APPLICATION NUMBER: US/09/902,535
; CURRENT FILING DATE: 2001-07-09
; PRIOR APPLICATION NUMBER: 60/216,657
; PRIOR FILING DATE: 2000-07-06
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 2930
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (930)...(2546)
US-09-902-535-1
Query Match 100.0%; Score 201; DB 10; Length 2930;
Best Local Similarity 100.0%; Pred. No. 6; le-54;
Matches 201; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 AACTTCAGCAGAGAAATTAACACACCTCTTGTAGAGACCTCTGTTCCTCAATCTG 60
|||||
DB 1551 AACTTCAGCAGAGAAATTAACACACCTCTTGTAGAGACCTCTGTTCCTCAATCTG 1610
OY 61 GAATTAACCATTAACCTCAACCTCACTGTTAAATTTGATCACTACTACATACACACC 120
|||||
DB 1611 GAATTAACCATTAACCTCAACCTCACTGTTAAATTTGATCACTACTACATACACACC 1670
OY 121 AACTCCCAATGATCAGTGGTGAATCTCTCCACACAATATGCTGCTACCTCAGGA 180
|||||
DB 1671 AACTCCCAATGATCAGTGGTGAATCTCTCTCCACACAATATGCTGCTACCTCAGGA 1730

QY 181 ATATTTTGTCTGTGTACC 201
|||||
DB 1731 ATATTTTGTCTGTGTACC 1751

RESULT 2
US-10-114-893-134

; Sequence 134, Application US/10114893
; Publication No. US20020193567A1

GENERAL INFORMATION:

APPLICANT: Jacobs, Kenneth
APPLICANT: McCoy, John M.
APPLICANT: Lavalley, Edward R.
APPLICANT: Collins-Racie, Lisa A.
APPLICANT: Evans, Cheryl
APPLICANT: Merberg, David
APPLICANT: Treacy, Maurice
APPLICANT: Bowman, Michael R.
APPLICANT: Spaulding, Vikki
APPLICANT: Carlin-Duckett, McKeough
APPLICANT: Kelleher, Kerry S.
APPLICANT: Genetics Institute, Inc.
TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES ENCODING THEM
FILE REFERENCE: GI 6000-10A
CURRENT APPLICATION NUMBER: US/10/114,893
CURRENT FILING DATE: 2002-04-02
EARLIER APPLICATION NUMBER: 09/413,232
EARLIER FILING DATE: 1999-10-06
NUMBER OF SEQ ID NOS: 321
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 134
LENGTH: 2946
TYPE: DNA
ORGANISM: Homo sapiens
US-10-114-893-134

Query Match 100.0%; Score 201; DB 9; Length 2946;
Best Local Similarity 100.0%; Pred. No. 6,2e-54;
Matches 201; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AACTTCAGACAGAAATAACACCACTCCGTTTAGTAGAGACCTCTGTTCACATCG 60
|||||
DB 1549 AACTTCAGACAGAAATAACACCACTCCGTTTAGTAGAGACCTCTGTTCACATCG 1608
QY 61 GAATAAACCACATACCTCAACCTCACTGTGTAAATTAGCAATACATACACAACC 120
|||||
DB 1609 GAATAAACCACATACCTCAACCTCACTGTGTAAATTAGCAATACATACACAACC 1668
QY 121 AACTCCCAATGATGAGTGGGTAACCTCTCCACACAATAAGTGTGCTTACCTCAGGA 180
|||||
DB 1669 AACTCCCAATGATGAGTGGGTAACCTCTCCACACAATAAGTGTGCTTACCTCAGGA 1728
QY 181 ATATTTTGTCTGTGTACC 201
|||||
DB 1729 ATATTTTGTCTGTGTACC 1749

RESULT 3

US-08-979-847-105

; Sequence 105, Application US/08979847
; Publication No. US20030039664A1

GENERAL INFORMATION:

APPLICANT: PERRON, HERVE
APPLICANT: BESEME, FREDERIC
APPLICANT: BEDIN, FREDERIC
APPLICANT: PARANHOS-BACCALA, GLAUCIA
APPLICANT: KOMURIAN-PRADEL, FLORENCE
APPLICANT: JOLIVET-REYNAUD, COLETTE
APPLICANT: MANDRAND, BERNARD
APPLICANT: GARSON, JEREMY
APPLICANT: TUKE, PHILIP
TITLE OF INVENTION: VIRAL MATERIAL AND NUCLEOTIDE FRAGMENTS

TITLE OF INVENTION: ASSOCIATED WITH MULTIPLE SCLEROSIS, FOR DIAGNOSTIC, PROPHYL
TITLE OF INVENTION: THERAPEUTIC PURPOSES
NUMBER OF SEQUENCES: 210
CORRESPONDENCE ADDRESS:
ADDRESSEE: OLIVIER & BERRIDGE, PLC
STREET: P.O. BOX 19928
CITY: ALEXANDRIA
STATE: VA
COUNTRY: USA
ZIP: 22320

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/979,847
FILING DATE: 26-NOV-1997
CLASSIFICATION: 4.35
ATTORNEY/AGENT INFORMATION:
NAME: BERRIDGE, WILLIAM P.
REGISTRATION NUMBER: 30,024
REFERENCE/DOCKET NUMBER: WPB 39046A
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-836-6400
TELEFAX: 703-836-2787
INFORMATION FOR SEQ ID NO: 105:
SEQUENCE CHARACTERISTICS:
LENGTH: 1481 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
US-08-979-847-105

Query Match 96.0%; Score 193; DB 7; Length 1481;
Best Local Similarity 97.5%; Pred. No. 1.6e-51;
Matches 196; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 AACTTCAGACAGAAATAACACCACTCCGTTTAGTAGAGACCTCTGTTCACATCG 60
|||||
DB 622 AACTTCAGACAGAAATAACACCACTCCGTTTAGTAGAGACCTCTGTTCACATCG 681
QY 61 GAATAAACCACATACCTCAACCTCACTGTGTAAATTAGCAATACATACACAACC 120
|||||
DB 682 GAATAAACCACATACCTCAACCTCACTGTGTAAATTAGCAATACATACACAACC 741
QY 121 AACTCCCAATGATGAGTGGGTAACCTCTCCACACAATAAGTGTGCTTACCTCAGGA 180
|||||
DB 742 AACTCCCAATGATGAGTGGGTAACCTCTCCACACAATAAGTGTGCTTACCTCAGGA 801
QY 181 ATATTTTGTCTGTGTACC 201
|||||
DB 802 ATATTTTGTCTGTGTACC 822

RESULT 4

US-09-864-761-21192

; Sequence 21192, Application US/09864761
; Patent No. US20020048763A1

GENERAL INFORMATION:

APPLICANT: Penn, Sharon G.
APPLICANT: Rank, David R.
APPLICANT: Hanzel, David K.
APPLICANT: Chen, Wensheng
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FO
FILE REFERENCE: Aecm1ca-x-1
CURRENT APPLICATION NUMBER: US/09/864,761
CURRENT FILING DATE: 2001-05-23
PRIOR APPLICATION NUMBER: US 60/180,312
PRIOR FILING DATE: 2000-02-04
PRIOR APPLICATION NUMBER: US 60/207,456

```

; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 21192
; LENGTH: 792
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC002346.1
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 20
; OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 6
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 5.4
; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 5.9
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 4
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 6.7
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 6.2
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 6.3
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 3.9
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 7.2
; OTHER INFORMATION: NT HIT: AF208161.1, EVALU0 0.00e+00
; OTHER INFORMATION: EST_HUMAN HIT: A0138405.1, EVALU0 0.00e+00
; OTHER INFORMATION: SWISSPROT HIT: P10269, EVALU0 2.00e-16
US-09-864-761-21192

Query Match      87.3%; Score 175.4; DB 10; Length 792;
Best Local Similarity 95.5%; Pred. No. 5e-46;
Matches 192; Conservative 0; Mismatches 6; Indels 3; Gaps 1;

QY 1 AACCTGACACGAATAAACCACCTGCTGTTAGTAGGACCTGTTGTTCCATCNG 60
DB 248 AACCTGACACGAATAAACCACCTGCTGTTAGTAGGCTCCTC---TTTCCAACTCG 304
QY 61 GAAATACCCATACCTCAACCTCACCTGTTGTAATTAAGCAATACATACACAGC 120
DB 305 GAAATACCCATACCTCAACCTCACCTGTTGTAATTAAGCAATACATACAGC 364
QY 121 AACCTGACATGATGAGTGGGTAAGTCTCTCCACACAAATAGTGCCTACCTAGGA 180
DB 365 AACCTGACATGATGAGTGGGTAAGTCTCTCCACACAAATAGTGCCTACCTAGGA 424
QY 181 AATATTTTGTCTGTGTGTACC 201
```

```

DB 425 AATATTTTGTCTGTGTGTACC 445
|||||
RESULT 5
US-09-864-761-4444
; Sequence 4444, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FO
; FILE REFERENCE: Aecomica-X-1
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 4444
; LENGTH: 1894
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC002346.1
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 20
; OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 6
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 5.4
; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 5.9
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 4
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 6.7
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 6.2
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 6.3
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 3.9
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 7.2
US-09-864-761-4444
```

Query Match	87.3%;	Score 175.4;	DB 10;	Length 1894;
Best Local Similarity	95.5%;	Pred. No. 7.1e-46;		
Matches 192; Conservative	0;	Mismatches 6;	Indels 3;	Gaps 11

QY	1	ACCTTCAGCAGAGAAATTAACACACAGCTCCGTTTATAGAGACCTCTGTTTCCAAATCTG	60
Db	446	AACCTTCAGCAGAGAAATTAACACACACAGCTCTTCTGTTTATAGAGTCTCTC--TTTCCAAATCTG	5028
QY	61	GAATAAACCCATACCTCAAAACCTCACCTCGTGTAAATTAATTAGCAATACATACACACACC	120
Db	503	GAATTAATACCCATACCTCAAAACCTCACCTCGTGTAAATTAATTAGCAATACATACACACAGCC	5628
QY	121	AACCTCCCAATGCATACAGGTGGGTACTCTCTCCACACACAATAGTCTGCTTACCTCAGGA	180
Db	563	AACCTCCCAATGCATACAGGTGGGTACTCTCTCCACACACAATAGTCTGCTTACCTCAGGA	6228
QY	181	ATATTTTTTGTCTGTGGTACC	201
Db	623	ATATTTTTTGTCTGTGGTACC	643

RESULT 6

```

US-09-880-107-2185/c
Sequence 2185, Application US/09880107
Patent No. US20020142981A1
GENERAL INFORMATION:
APPLICANT: Horne, Darci T.
APPLICANT: Vockley, Joseph G.
APPLICANT: Scherr, Uwe
TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer
FILE REFERENCE: 44921-5028-WO
CURRENT APPLICATION NUMBER: US/09/880,107
CURRENT FILING DATE: 2001-06-14
PRIOR APPLICATION NUMBER: US 60/211,379
PRIOR FILING DATE: 2000-06-14
PRIOR APPLICATION NUMBER: US 60/237,054
PRIOR FILING DATE: 2000-10-02
NUMBER OF SEQ ID NOS: 3950
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 2185
LENGTH: 3457
Type: DNA
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: Genbank Accession No. US20020142981A1 L26336
US-09-880-107-2185

```

Query Match	16.6%;	Score 33.4;	DB 10;	Length 3457;
Best Local Similarity	52.5%;	Pred. No. 1.4;		
Matches 73; Conservative	0;	Mismatches 66;	Indels 0;	Gaps 0;

QY	7	AGCGCAAAATTAATTAACACCACTTCGCTTTAGTAGAGACCTCTTGTTCCATCTGGAATA	66
Db	3370	AGTACATAAAAAATTAACCAATTTAGTCCGTATATATATTTACGTGTTCCAGAGTTCTACAG	3311
QY	67	ACCATTACCTCAACCTCCACCTGTGTAAATTTAGCAATATACATATACACAACCACTCC	126
Db	3310	AACAGGTGCCAAACTTAAGGTTTCTCATCTCAACCAATCTCATGACGAACTAATCA	3251
QY	127	CAATGCATCAGGTGGGTAA	145
Db	3250	AAAAAGCTCAAGTCAGCA	3222

RESULT 7

US-09-938-842A-3786/c
Sequence 3786, Application US/09938842A
Patent No. US20020160378A1
GENERAL INFORMATION:
APPLICANT: Harper, Jeff
APPLICANT: Kreps, Joel

```

1  APPLICANT: Wang, Xun
2  APPLICANT: Zhu, Tong
3  TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
4  TITLE OF INVENTION: SAME, AND METHODS OF USE
5  FILE REFERENCE: SCRIPT300-3
6  CURRENT APPLICATION NUMBER: US/09/938, 842A
7  CURRENT FILING DATE: 2001-08-24
8  PRIOR APPLICATION NUMBER: US 60/227, 866
9  PRIOR FILING DATE: 2000-08-24
10 PRIOR APPLICATION NUMBER: US 60/264, 647
11 PRIOR FILING DATE: 2001-01-16
12 PRIOR APPLICATION NUMBER: US 60/300, 111
13 PRIOR FILING DATE: 2001-06-22
14 NUMBER OF SEQ ID NOS: 5379
15 SEQ ID NO 3786
16 LENGTH: 1084
17 TYPE: DNA
18 ORGANISM: Arabidopsis thaliana
19 US-09-938-842A-3786

```

Query Match	16.1%;	Score 32.4;	DB 9;	Length 1084;
Best Local Similarity	56.6%;	Pred. No. 1.8;		
Matches 60;	Conservative 0;	Mismatches 46;	Indels 0;	Gaps 0;

Qy	Db	Qy	Db
11	576	71	516
CAGAAATTAACACCACTCCGTTTATAGAGACCTCTGTGTCAAATCGAATTAACCC	CATAAATTAACAAACGATCTGTACTCTTTGTAATTTTAATCGTATCTATACCTCAAA	ATACCTCAACCTCACCTGTGTAAATTTAGCAATCTACTACTACAC	AAAAAGTAAACAAACAGTTACAAATTAATCATGAATACATATAC
70	517	116	471

RESULT 8

```
US-09-887-576-143/C
Sequence 143: Application US/09867576
Patent No. US20020144047A1
GENERAL INFORMATION:
APPLICANT: Budworth, P.
APPLICANT: Brown, D.
APPLICANT: Chang, H.
APPLICANT: Zhu, T.
APPLICANT: Han, B.
APPLICANT: Wang, X.
APPLICANT: Cooper, Bret
TITLE OF INVENTION: Promoters for regulation of plant expression
FILE REFERENCE: 1360.001US1
CURRENT APPLICATION NUMBER: US/09/887,576
CURRENT FILING DATE: 2001-06-25
PRIOR APPLICATION NUMBER: US 60/213,848
PRIOR FILING DATE: 2000-06-23
PRIOR APPLICATION NUMBER: US 60/214,087
PRIOR FILING DATE: 2000-06-23
PRIOR APPLICATION NUMBER: US 60/258,692
PRIOR FILING DATE: 2000-12-29
NUMBER OF SEQ ID NOS: 875
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 143
LENGTH: 2000
TYPE: DNA
ORGANISM: Arabidopsis thaliana
US-09-887-576-143
```

Query Match	16.1%;	Score 32.4;	DB 10;	Length 2000;
Best Local Similarity	56.6%;	Pred. No. 2.3;		
Matches 60;	Conservative 0;	Mismatches 46;	Indels 0;	Gaps 0;

QY 11 CAGAAATAAACACCCTCTCGTTTGTATGAGACCTCTGTTTCCAATCTGGAAATAACCC 70
| | | | | | | | | | | | | | | | | | | | | |
Db 1491 CATTAATAAAGAACAATTGATAGTCTTGTAATTTTAATCATCTATCTACTCCAAATACAA 1432
| | | | | | | | | | | | | | | | | | | | | |
QY 71 ATACCTCAAACCTCACCTGTGTAATAATTAGCAATATCTATCATACAC 116
| | | | | | | | | | | | | | | | | | | | | |

Db 1431 AAAACGTAACAAACAGTTACAAATTAATCATGATACAAATCAC 1386

RESULT 9

US-09-738-626-1
; Sequence 1, Application US/09738626
; Publication No. US20020197605A1
; GENERAL INFORMATION:
; APPLICANT: NAKAGAWA, SATOSHI
; APPLICANT: MIZOGUCHI, HIROSHI
; APPLICANT: ANDO, SEIKO
; APPLICANT: HAYASHI, MIKIRO
; APPLICANT: OCHIAI, KEIKO
; APPLICANT: YOKOI, HARUHIKO
; APPLICANT: TATEISHI, NAOKO
; APPLICANT: SENO, AKIHIRO
; APPLICANT: IKEDA, MASATO
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-125
; CURRENT APPLICATION NUMBER: US/09/738,626
; PRIOR FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: JP 99/377484
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: JP 00/159162
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: JP 00/280988
; PRIOR FILING DATE: 2000-08-03
; NUMBER OF SEQ ID NOS: 7059
; SOFTWARE: PatentIn ver. 3.0
; SEQ ID NO 1
; LENGTH: 3309400
; TYPE: DNA
; ORGANISM: Corynebacterium glutamicum
US-09-738-626-1

Query Match

Best Local Similarity 15.2%; Score 30.6; DB 9; Length 3309400;
Matches 54; Conservative 0; Mismatches 39; Indels 0; Gaps 0;

QY 92 TAAATTTAGCATCTACTACACACCAATCCCAATGCAATGAGTGGTAACTCCTC 151

Db 1913847 TTTAAATTTACAAATTAATAACAAACATACACTGATTTAACTCAACTAACGCCCC 1913906

QY 152 CCACACAAATAGTCTGCTACCCCTCAGAGAAAT 184

Db 1913907 CAAACATACAAAAAGCACACCTCTCAGACATCT 1913939

RESULT 10

US-09-902-535-5
; Sequence 5, Application US/09902535
; Patent No. US20020102530A1
; GENERAL INFORMATION:
; APPLICANT: Keith, Jr., James C.
; APPLICANT: McCoy, John M.
; APPLICANT: MI, Sha
; TITLE OF INVENTION: Methods and compositions for diagnosing
; TITLE OF INVENTION: and treating preclampsia and gestational trophoblast
; FILE REFERENCE: GIN-6006B4
; CURRENT APPLICATION NUMBER: US/09/902,535
; CURRENT FILING DATE: 2001-07-09
; PRIOR APPLICATION NUMBER: 60/216,657
; PRIOR FILING DATE: 2000-07-06
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 30
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-902-535-5

Query Match 14.9%; Score 30; DB 10; Length 30;
Best Local Similarity 100.0%; Pred. No. 2.3;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 146 CTCCTCCACACAAATGTCTGCTACCT 175

Db 1 CTCCTCCACACAAATGTCTGCTACCT 30

RESULT 11

US-09-822-830A-567/c
; Sequence 567, Application US/09822830A
; Patent No. US20020142952A1
; GENERAL INFORMATION:
; APPLICANT: Genetics Institute, Inc.
; APPLICANT: Wong, Gordon G.
; APPLICANT: Clark, Hilary
; APPLICANT: Fehchel, Kim
; APPLICANT: Agostino, Michael J.
; APPLICANT: Howes, Steven H.
; APPLICANT: Resnick, Richard J.
; APPLICANT: Gulukota, Kamalakara
; APPLICANT: Graham, James R.
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING NOVEL SECRETED PROTEINS
; FILE REFERENCE: GIN 6402
; CURRENT APPLICATION NUMBER: US/09/822,830A
; PRIOR FILING DATE: 2001-03-29
; PRIOR APPLICATION NUMBER: 60/195,604
; PRIOR FILING DATE: 2000-04-06
; NUMBER OF SEQ ID NOS: 631
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 567
; LENGTH: 580
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-822-830A-567

Query Match

Best Local Similarity 14.9%; Score 30; DB 10; Length 580;
Matches 51; Conservative 0; Mismatches 35; Indels 0; Gaps 0;

QY 33 TTTAGTAGACCTCTGTTTCCATCTGGAATTAACCAATCAACCTGCTGTGT 92

Db 576 TTTAGACACAGACTTGTATCTATATAGATATAATCATCTCAATTAACAAATCTAAT 517

QY 93 AAAATTTAGCAATCTACATACACAA 118

Db 516 AATATTTACAAATTAATTTATACACAA 491

RESULT 12

US-09-822-830A-123/c
; Sequence 123, Application US/09822830A
; Patent No. US20020142952A1
; GENERAL INFORMATION:
; APPLICANT: Genetics Institute, Inc.
; APPLICANT: Wong, Gordon G.
; APPLICANT: Clark, Hilary
; APPLICANT: Fehchel, Kim
; APPLICANT: Agostino, Michael J.
; APPLICANT: Howes, Steven H.
; APPLICANT: Resnick, Richard J.
; APPLICANT: Gulukota, Kamalakara
; APPLICANT: Graham, James R.
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING NOVEL SECRETED PROTEINS
; FILE REFERENCE: GIN 6402
; CURRENT APPLICATION NUMBER: US/09/822,830A
; CURRENT FILING DATE: 2001-03-29
; PRIOR APPLICATION NUMBER: 60/195,604
; PRIOR FILING DATE: 2000-04-06
; NUMBER OF SEQ ID NOS: 631
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 123

LENGTH: 1962
TYPE: DNA
ORGANISM: Homo sapiens
US-09-822-830A-123

Query Match 14.9% Score 30; DB 10; Length 1962;
Best Local Similarity 59.3%; Pred. No. 13;
Matches 51; Conservative 0; Mismatches 35; Indels 0; Gaps 0;

QY 33 TTATAGAGACCTCTGTTCATCTGGAATTAACCACTTACTTCAACCTGCTGT 92
DB 1202 TTATAGACAGACTTGTATCTATATTAATCAATGCTTAATCAAAATCTAAAT 1143

QY 93 AAAATTAGCACTACTACATACAA 118
DB 1142 AATATTACAATAATATTACACA 1117

RESULT 13

US-09-925-301-15/C
Sequence 15, Application US/09925301
Patent No. US20020052308A1
GENERAL INFORMATION:

APPLICANT: Rosen et al.
TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
FILE REFERENCE: PA106
CURRENT APPLICATION NUMBER: US/09/925,301
CURRENT FILING DATE: 2001-08-10
PRIOR APPLICATION NUMBER: PCT/US00/05882
PRIOR FILING DATE: 2000-03-08
PRIOR APPLICATION NUMBER: 60/124,270
PRIOR FILING DATE: 1999-03-12
NUMBER OF SEQ ID NOS: 1694
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 15

LENGTH: 1024

TYPE: DNA

ORGANISM: Homo sapiens

FEATURE:

NAME/KEY: misc_feature

LOCATION: (938)

OTHER INFORMATION: n equals a,t,g, or c

NAME/KEY: misc_feature

LOCATION: (1005)

OTHER INFORMATION: n equals a,t,g, or c

NAME/KEY: misc_feature

LOCATION: (1012)

OTHER INFORMATION: n equals a,t,g, or c

NAME/KEY: misc_feature

LOCATION: (1019)

OTHER INFORMATION: n equals a,t,g, or c

US-09-925-301-15

Query Match 14.6% Score 29.4; DB 10; Length 1024;
Best Local Similarity 50.8%; Pred. No. 16;
Matches 60; Conservative 3; Mismatches 55; Indels 0; Gaps 0;

QY 84 CACCGTGTGAATTTAGCAATCTACATACACCAACTCCCATGATAGTGGGT 143
DB 964 CACCGTACTCGACATGTGACAGTANCAAGCAATAGCTTGACTGTGCAYCATGAGGT 905

QY 144 AACCTCCACCAAAATAGTGTGCTACCTCAGGAATATTTTGTGTGTACC 201
DB 904 TTCTCTCTCCCTACTTAATATTCATCAATCTTCAGCAGCTTTTATCAACACTGTTCC 847

RESULT 14

US-08-781-986A-289

Sequence 289, Application US/08781986A

Publication No. US20030054436A1

GENERAL INFORMATION:

APPLICANT: Charles Kunsch

TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences

NUMBER OF SEQUENCES: 5255

CORRESPONDENCE ADDRESS:

ADDRESSEE: Human Genome Sciences, Inc.

STREET: 9410 Key West Avenue

CITY: Rockville

STATE: Maryland

COUNTRY: USA

ZIP: 20850

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette, 3.50 inch, 1.4mb storage

COMPUTER: HP Vectra 486/33

OPERATING SYSTEM: MSDOS version 6.2

SOFTWARE: ASCII Text

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/781,986A

FILING DATE:

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER:

FILING DATE:

ATTORNEY/AGENT INFORMATION:

NAME: Benson, Bob

REGISTRATION NUMBER: 30,446

REFERENCE/DOCKET NUMBER: P8248PP

TELECOMMUNICATION INFORMATION:

TELEPHONE: (301) 309-8504

TELEFAX: (301) 309-8512

INFORMATION FOR SEQ ID NO: 289:

SEQUENCE CHARACTERISTICS:

LENGTH: 1999 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

US-08-781-986A-289

Query Match 14.6% Score 29.4; DB 7; Length 1999;
Best Local Similarity 54.1%; Pred. No. 21;
Matches 60; Conservative 0; Mismatches 51; Indels 0; Gaps 0;

QY 5 TCAGCAGCAATTAACACCACTTCCGTTTGTAGTAGGACCTTGTTCATCTCGAAA 64
DB 638 TTACCATGAACTATATATCTTATGCTCAAAATATATCTTGGCAGGAAATCCAGAA 697

QY 65 TAACCATCTCAACCTGACCTGCTGTAATTTAGCAATCTACATACA 115
DB 698 TATCATTTATCTTACCTCTACATTTTATATTTAAGTAATCTCCATACA 748

RESULT 15

US-09-263-959-1

Sequence 1, Application US/09263959

Patent No. US20020150891A1

GENERAL INFORMATION:

APPLICANT: Hood, Leroy E.

APPLICANT: Bowen, Lee

APPLICANT: Koop, Ben F.

TITLE OF INVENTION: DIAGNOSTIC AND THERAPEUTIC COMPOSITIONS AND METHODS WHICH U

CORRESPONDENCE ADDRESS:

ADDRESSEE: Seed and Berry LLP

STREET: 6300 Columbia Center, 701 Fifth Avenue

CITY: Seattle

STATE: Washington

COUNTRY: US

ZIP: 98104-7092

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/263,959

FILING DATE: 05-MAR-1999

GenCore version 5.1.5
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 2, 2003, 12:54:45 ; Search time 689 Seconds

(Without alignments)
8490.078 Million cell updates/sec

Title: US-09-719-554-3_COPY_9000_9200

Perfect score: 201
Sequence: 1 cctagagagagtagtgcctc.....actgagacctgggacctct 201

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

```
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vi:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_mus:*
34: em_htg_pln:*
35: em_htg_rtd:*
36: em_htg_mam:*
37: em_htg_vrt:*
38: em_sy:*
39: em_htgo_hum:*
40: em_htgo_mus:*
41: em_htgo_other:*
```

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	201	100.0	2055	6	AX007999	AX007999 Sequence
2	201	100.0	2264	6	HSRA57	X33499 H.sapiens m
3	201	100.0	2599	6	AX007978	AX007978 Sequence
4	201	100.0	2781	6	AX092223	AX092223 Sequence
5	201	100.0	2781	9	AF072506	AF072506 Homo sapi
6	201	100.0	2930	6	AX355872	AX355872 Sequence
7	201	100.0	2930	6	AF208161	AF208161 Homo sapi
8	201	100.0	2946	9	AK177269	AK177269 Sequence
9	201	100.0	10499	6	AX007980	AX007980 Sequence
10	201	100.0	56093	6	AX329572	AX329572 Sequence
11	201	100.0	56093	9	HSAC000064	AC000064 Human BAC
12	201	100.0	149194	9	AC007566	AC007566 Homo sapi
13	199.4	99.2	1617	9	AF156963	AF156963 Homo sapi
14	199.4	99.2	1624	9	AF506835	AF506835 Homo sapi
15	199.4	99.2	1860	9	AF513360	AF513360 Homo sapi
16	198.2	98.6	7582	6	AX000966	AX000966 Sequence
17	198.2	98.6	7582	6	AX027480	AX027480 Sequence
18	197.8	98.4	1136	6	AX000961	AX000961 Sequence
19	197.8	98.4	1136	6	AX027475	AX027475 Sequence
20	197.8	98.4	1136	9	AF072508	AF072508 Homo sapi
21	193	96.0	2782	6	AX000962	AX000962 Sequence
22	193	96.0	2782	6	AX027476	AX027476 Sequence
23	193	96.0	2782	6	AX133396	AX133396 Sequence
24	188.2	93.6	172281	6	AC068492	AC068492 Homo sapi
25	186.6	92.8	108439	2	AC019325	AC019325 Homo sapi
26	186.6	92.8	125790	9	AC106713	AC106713 Homo sapi
27	186.6	92.8	176087	2	AC024625	AC024625 Homo sapi
28	186.6	92.8	176459	2	AC108684	AC108684 Homo sapi
29	186.6	92.8	178333	9	AC108685	AC108685 Homo sapi
30	185	92.0	1684	6	AX400057	AX400057 Sequence
31	185	92.0	2074	6	AX478550	AX478550 Sequence
32	185	92.0	104853	9	AC117444	AC117444 Homo sapi
33	185	92.0	144900	9	AL391219	AL391219 Human DNA
34	185	92.0	162579	9	AL390039	AL390039 Human DNA
35	185	92.0	169462	2	AC024033	AC024033 Homo sapi
36	185	92.0	172918	2	AC023366	AC023366 Homo sapi
37	185	92.0	176305	2	AC074086	AC074086 Homo sapi
38	185	92.0	186723	2	AC027752	AC027752 Homo sapi
39	185	92.0	187321	2	AC092510	AC092510 Papio cyn
40	183.4	91.2	591	6	AX000986	AX000986 Sequence
41	183.4	91.2	591	6	AX027469	AX027469 Sequence
42	183.4	91.2	591	9	AF072498	AF072498 Homo sapi
43	183.4	91.2	1481	6	AX001027	AX001027 Sequence
44	183.4	91.2	1629	6	AX036992	AX036992 Sequence
45	183.4	91.2	1629	14	AF331500	AF331500 Multiple

ALIGNMENTS

```
RESULT 1
LOCUS AX007999 2055 bp DNA linear PAT 06-SEP-2000
DEFINITION Sequence 22 from Patent WO9967395.
ACCESSION AX007999
VERSION AX007999.1 GI:9995696
KEYWORDS
SOURCE
ORGANISM human.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1 (bases 1 to 2055)
Perlin,J.P., Rieger,F. and Alliel,P.M.
Nucleic sequence and deduced protein sequence family with human
endogenous retroviral motifs, and their uses
```

JOURNAL Patent: WO 9667395-A 22-29-DEC-1999;
INST NAT SANTE RECH MED (FR); PERIN JEAN PIERRE (FR); RIEGER
FRANCOIS (FR); ALBIEU PATRICK M (FR)
FEATURES location/Qualifiers
source 1. 2055
/organism="Homo sapiens"
/db_xref="taxon:9606"
CDS <1. 87
/note="unnamed protein product"
/codon_start=1
/protein_id="CA007392.1"
/db_xref="GI:9995697"
/translation="PKTANLVADITSLAKYQVLTLOGTYP"
BASE COUNT 576 a 574 c 376 g 529 t
ORIGIN

Query Match 100.0%; Score 201; DB 6; Length 2055;
Best Local Similarity 100.0%; Pred. No. 1.7e-48;
Matches 201; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCTAGCAGCAGTAGTCTCTCAAAATGCAAGAGCTTGAAGCTTCTAACCCTGAAGAGG 60
|||
Db 1512 CCTAGCAGCAGTAGTCTCTCAAAATGCAAGAGCTTGAAGCTTCTAACCCTGAAGAGG 1571
|||
QY 61 GGGACCTGTTATTTTATTTAGGGGAGAGATGCTGTATTTATTTAATCAATCCGAATCGT 120
|||
Db 1572 GGGACCTGTTATTTTATTTAGGGGAGAGATGCTGTATTTATTTAATCAATCCGAATCGT 1631
|||
QY 121 CACTGAGAAAGTTAAGAATTCGAGATGCAATACAGCAGAGAGAGAGCTTGAAA 180
|||
Db 1632 CACTGAGAAAGTTAAGAATTCGAGATGCAATACAGCAGAGAGAGAGCTTGAAA 1691
|||
QY 181 CACTGAGACCTGGGGCTCTCT 201
|||
Db 1692 CACTGAGACCTGGGGCTCTCT 1712
|||

RESULT 2
HSRAB7 2264 bp mRNA linear PRI 10-FEB-1997
LOCUS HSRAB7
DEFINITION H.sapiens mRNA for Rab7 protein.
ACCESSION X93499
VERSION X93499.1 GI:1089892
KEYWORDS GTP-binding protein; rab7 gene; Rab7 protein.
SOURCE Homo sapiens.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 2264)
Vitel, R., Chiarillo, M., Lattero, D., Bruni, C.B. and Buccì, C.
Molecular cloning and expression analysis of the human Rab7 GTP-ase
complementary deoxyribonucleic acid
Biochem. Biophys. Res. Commun. 229 (3), 887-890 (1996)
JOURNAL
MEDLINE 97115674
PUBMED 8954989
REFERENCE 2 (bases 1 to 2264)
Bucci, C.
Direct Submission
Submitted (27-NOV-1995) C. Bucci, University of Napoli, Dipt di
Biologia e Patologia Cellulare, e Molecolare, II Policlinico, via
Pansini 5, I-80131 Napoli, ITALY
FEATURES location/Qualifiers
source 1. 2264
/organism="Homo sapiens"
/db_xref="taxon:9606"
/tissue_type="placenta"
603. 1226
/gene="Rab7"
603. 1226
/gene="Rab7"
/function="GTP-binding protein"
/codon_start=1
/product="Rab7 protein"

/protein_id="CA063763.1"
/db_xref="GI:1089893"
/db_xref="SWISS-PROT:P51149"
/translation="MTSRKVLKVIILGDSVGKTSIMNOYVKNKFSNOYKATIGAD
FLTKREVMDDRLVTMOIMDTAGCEERFOSGVFAYGADCCVLPVTPANPEKTDLSM
KDEFLIQASPRPEPFVYLGNKIDLEKROYATRAQAMCKSKNNIPFEFSAREAI
NVEQAFQTLARNALKEFEVELNEPEPIKIDNDRAKASAESCS"
BASE COUNT 623 a 600 c 551 g 490 t
ORIGIN

Query Match 100.0%; Score 201; DB 9; Length 2264;
Best Local Similarity 100.0%; Pred. No. 1.8e-48;
Matches 201; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCTAGCAGCAGTAGTCTCTCAAAATGCAAGAGCTTGAAGCTTCTAACCCTGAAGAGG 60
|||
Db 1848 CCTAGCAGCAGTAGTCTCTCAAAATGCAAGAGCTTGAAGCTTCTAACCCTGAAGAGG 1907
|||
QY 61 GGGACCTGTTATTTTATTTAGGGGAGAGATGCTGTATTTATTTAATCAATCCGAATCGT 120
|||
Db 1908 GGGACCTGTTATTTTATTTAGGGGAGAGATGCTGTATTTATTTAATCAATCCGAATCGT 1967
|||
QY 121 CACTGAGAAAGTTAAGAATTCGAGATGCAATACAGCAGAGAGAGAGCTTGAAA 180
|||
Db 1968 CACTGAGAAAGTTAAGAATTCGAGATGCAATACAGCAGAGAGAGAGCTTGAAA 2027
|||
QY 181 CACTGAGACCTGGGGCTCTCT 201
|||
Db 2028 CACTGAGACCTGGGGCTCTCT 2048
|||

RESULT 3
AX007978 2599 bp DNA linear PAT 06-SEP-2000
LOCUS AX007978
DEFINITION Sequence 1 from Patent WO9967395.
ACCESSION AX007978
VERSION AX007978.1 GI:9995675
KEYWORDS human.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 2599)
Perin, J.P., Rieger, F. and Albieu, P.M.
Nucleic sequence and deduced protein sequence family with human
endogenous retroviral motifs, and their uses
Patent: WO 9667395-A 1 29-DEC-1999;
INST NAT SANTE RECH MED (FR); PERIN JEAN PIERRE (FR); RIEGER
FRANCOIS (FR); ALBIEU PATRICK M (FR)
FEATURES location/Qualifiers
source 1. 2599
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 744 a 718 c 495 g 642 t
ORIGIN

Query Match 100.0%; Score 201; DB 6; Length 2599;
Best Local Similarity 100.0%; Pred. No. 1.8e-48;
Matches 201; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCTAGCAGCAGTAGTCTCTCAAAATGCAAGAGCTTGAAGCTTCTAACCCTGAAGAGG 60
|||
Db 2050 CCTAGCAGCAGTAGTCTCTCAAAATGCAAGAGCTTGAAGCTTCTAACCCTGAAGAGG 2109
|||
QY 61 GGGACCTGTTATTTTATTTAGGGGAGAGATGCTGTATTTATTTAATCAATCCGAATCGT 120
|||
Db 2110 GGGACCTGTTATTTTATTTAGGGGAGAGATGCTGTATTTATTTAATCAATCCGAATCGT 2169
|||
QY 121 CACTGAGAAAGTTAAGAATTCGAGATGCAATACAGCAGAGAGAGAGCTTGAAA 180
|||
Db 2170 CACTGAGAAAGTTAAGAATTCGAGATGCAATACAGCAGAGAGAGAGCTTGAAA 2229
|||
QY 181 CACTGAGACCTGGGGCTCTCT 201
|||

[illegible]

REFERENCE	AUTHORS	TITLE	JOURNAL	REMARK	COMMENT	FEATURES
1	Blond, J.-L., Beseme, F. and Mallet, F.	Direct Submission	Submitted (19-JUN-1998)	UM103 CNRS-biomerieux, ENS Lyon, 46 allée d'Italie, Lyon, Cedex 07 69364, France	3 (bases 1 to 2781)	Bouton, O., Blond, J.-L. and Mallet, F.
2	Submitted (11-MAY-1999)	UM103 CNRS-biomerieux, ENS Lyon, 46 allée d'Italie, Lyon, Cedex 07 69364, France	Sequence update by submitter	On May 11, 1999 this sequence version replaced gi:4262291.	Location/Qualifiers	
3	1. .2781	/organism="Homo sapiens"	/db_xref="taxon:9606"	/clone="cl.PH74"	/tissue_type="placenta"	/clone_id="Clonotech 5'-stretch plus library, Cat number H50148"
4	1. .2781	/rpt_family="HERV-W"	<1. .49	/note="R"	50. .503	507. .524
5	507. .524	/note="putative primer binding site for tRNA-w"	534. .535	/note="putative splice junction"	762. .2378	/codon_start=-1
6	/product="envelope protein precursor"	/protein_id="AADI546.2"	/db_xref="GI:4773880"	/translation="MALPYHFLFETVLLPFTLTAPPPCRCMTSSSPYQEFIRMRORP	GNIDAPYRSLSKGTPEPTAHTHPRRCYSATLTCMANHYHTGKINSCCGGLCV	TVCNTRYTQTGMSDGGGVQDARKHKEYISQLTRVHGTSPPYKGLDLSKHETLET
7	HTRLVSLFNTLTGLGHEVSAQNPTNCWICLPLNRPVSIIPVEQNNFSTELNTTSSV	LVGLVSLNLEITHTLSNLTGKFSNTTYTNSOCIRWTPPTQIYCLPSGIEFVCGTSA	YRCINGTSESMCELSFVPPMTIYTEDDLXSYVSKPRNRPVLPVPIAGVYGLAGL	YTGIGITSTOIFYKLSOELNGDMERADSLVTLQDLSLAAYVLQNRALDLLTAE	RGCGLFLGEECCYVNOSSIVTEKKEIRDRIRAREELRNQPMGLLSQMPWILP	PLGLAAILILLRPGCTFNLVNFVSSRLAVAKLQMEPMQSKTKYLRPLDRPASP
8	RSVDNDIKGTPEEISNAOPLRPSAGSS"	762. .824	825. .2375	/product="envelope protein"	1701. .1712	/note="furin cleavage site"
9	1713. .2375	/note="transmembrane-region site"	1890. .1949	/note="Region: immunosuppressive region"	2347. .2763	/note="U3-R junction undetermined"
10	2425. .2435	/note="polypurine tract"	2745. .2750	2762. .770 c	556 g	708 t
11	BASE COUNT	747 a	770 c	556 g	708 t	
12	ORIGIN					
13	Query Match	100.0%	Score 201, DB 9, Length 2781,			
14	Best Local Similarity	100.0%	Pred. No.1.8e-48,			
15	Matches 201, Conservative	0, Mismatches	0, Indels	0, Gaps	0, Gaps	0, Gaps
16	1	CCGAGCAGCAGTACTCTTCAAAATCGAAGAGCTTTAGACTTGTCTAACCGCTGAAGAGG	60			
17	1883	CCCTGAGCAGCAGTACTCTTCAAAATCGAAGAGCTTTAGACTTGTCTAACCGCTGAAGAGG	1942			
18	61	GGGAACTGTTTATTTTATTTAGGGAAGAATGCTGTTATTTATGTTATCAATCCGAATCGT	120			
19	1943	GGGAACTGTTTATTTTATTTAGGGAAGAATGCTGTTATTTATGTTATCAATCCGAATCGT	2002			

QY 121 CACTGAGAAAGTTAAGAAATTCAGATGCAATACAGCTAGACGAGAGCTTCGAAA 180
|||||
Db 2003 CACTGAGAAAGTTAAGAAATTCAGATGCAATACAGCTAGACGAGAGCTTCGAAA 2062
QY 181 CACTGAGACCTGGGGCTCT 201
|||||
Db 2063 CACTGAGACCTGGGGCTCT 2083

RESULT 6
AX355872 2930 bp DNA linear PAT 06-FEB-2002
LOCUS AX355872
DEFINITION Sequence 1 from Patent WO0204678.
ACCESSION AX355872
VERSION AX355872.1 GI:18620523
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
AUTHORS Keith, J.C., McCoy, J.M. and M.S.
TITLE Methods and compositions for diagnosing and treating preclampsia
JOURNAL and gestational trophoblast disorders
GENETICS Patent: WO 0204678-A 1 17-JAN-2002;
INSTITUTE, INC. (US)
FEATURES
source location/Qualifiers
1..2930
/organism="Homo sapiens"
/db_xref="taxon:9606"
930..2546
/note="unnamed protein product"
/codon_start=1
/protein_id="CAD23000.1"
/db_xref="GI:18620524"
/translation="MALPPIHFLFTVLLPSEFTLLAPPCCMTSSSPYOFELRMORP
GNIDAPSYRSLKGTPTFAHTHMPKCYSAITLCMHANHYTGKINSCGGLAY
TVGCTVFTQMGSDGGGVDOAEKHYEISOLTRVHGTSPPKGLDLSKLTHT
HTLVLSFNTLLGLHVSANFTNCWICLPINRPRYSIPVEQNNSTELINTSV
LVGPLVSNLEITHHTSNLTCVKFSNTTYTNSOCIRWVTPPTQIVCLPSGIFVCGTSA
YRLGINSSESMCLSPFLVPMITLYTEODLYNYISKRNKRVPILPEVIGAVYGAAG
TGIGGTTSTQFYKLSOELNGEMERADSLVTLQDLSLAIVLQNRALDLTAE
RGCTCLFEGECYVYNOGSIYEKVEIRDRIOREAEELRTNGPGLLSQMPHILP
FLGPLAIIILLFEGPCIFNLVNFVSSRIEAVKLOMERPMOSKTKIYRPLDRPAS
RSDVNDIKGTPEEISNAOPLRPNAGSS"

BASE COUNT 842 a 800 c 571 g 717 t
ORIGIN

Query Match 100.0%; Score 201; DB 6; Length 2930;
Best Local Similarity 100.0%; Pred. No. 1.8e-48;
Matches 201; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCTAGCAGCAGTAGTCTCTCAAAATCGAAGAGCTTTAGACTTCTAACCGCTGAAGAGG 60
|||||
Db 2051 CCTAGCAGCAGTAGTCTCTCAAAATCGAAGAGCTTTAGACTTCTAACCGCTGAAGAGG 2110

QY 61 GGAACCTGTTATTTTATAGGGGAAGATGCTGTATTATGTTATATCCGAATCGT 120
|||||
Db 2111 GGAACCTGTTATTTTATAGGGGAAGATGCTGTATTATGTTATATCCGAATCGT 2170

QY 121 CACTGAGAAAGTTAAGAAATTCAGATGCAATACAGCTAGACGAGAGCTTCGAAA 180
|||||
Db 2171 CACTGAGAAAGTTAAGAAATTCAGATGCAATACAGCTAGACGAGAGCTTCGAAA 2230

QY 181 CACTGAGACCTGGGGCTCT 201
|||||
Db 2231 CACTGAGACCTGGGGCTCT 2251

RESULT 7
AF208161 2930 bp mRNA linear PRI 22-FEB-2000
LOCUS AF208161

DEFINITION Homo sapiens syncytin precursor, mRNA, complete cds.
ACCESSION AF208161
VERSION AF208161.1 GI:6760400
KEYWORDS
SOURCE Homo sapiens.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
AUTHORS Mi, S., Lee, X., Li, X., Veldman, G.M., Finnerty, H., Racie, L.,
Lavallie, E., Tang, X.Y., Edouard, P., Howes, S., Keith, J.C. Jr. and
McCoy, J.M.
TITLE Syncytin is a captive retroviral envelope protein involved in human
JOURNAL placental morphogenesis
MEDLINE Nature 403 (6771), 785-789 (2000)
PUBMED 20155476
10693809

REFERENCE
AUTHORS 2 (bases 1 to 2930)
She, M., Lee, X., Li, X., Veldman, G.M., Finnerty, H., Racie, L.,
Lavallie, E., Tang, X., Edouard, P., Howes, S., Keith, J.C. Jr. and
McCoy, J.M.
TITLE Direct Submision
JOURNAL Submitted (26-NOV-1999) Genetics Institute, 87 Cambridge Park
Drive, Cambridge, MA 02140, USA
FEATURES
source location/Qualifiers
1..2930
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="AJ172-25"
/tissue_type="testis"
1..2930
/rpl_family="Human endogeneous retrovirus HERV-W"
1..929
930..2546
/note="envlope protein"
/codon_start=1
/product="syncytin precursor"
/protein_id="AAE28334.1"
/db_xref="GI:6760401"
/translation="MALPPIHFLFTVLLPSEFTLLAPPCCMTSSSPYOFELRMORP
GNIDAPSYRSLKGTPTFAHTHMPKCYSAITLCMHANHYTGKINSCGGLAY
TVGCTVFTQMGSDGGGVDOAEKHYEISOLTRVHGTSPPKGLDLSKLTHT
HTLVLSFNTLLGLHVSANFTNCWICLPINRPRYSIPVEQNNSTELINTSV
LVGPLVSNLEITHHTSNLTCVKFSNTTYTNSOCIRWVTPPTQIVCLPSGIFVCGTSA
YRLGINSSESMCLSPFLVPMITLYTEODLYNYISKRNKRVPILPEVIGAVYGAAG
TGIGGTTSTQFYKLSOELNGEMERADSLVTLQDLSLAIVLQNRALDLTAE
RGCTCLFEGECYVYNOGSIYEKVEIRDRIOREAEELRTNGPGLLSQMPHILP
FLGPLAIIILLFEGPCIFNLVNFVSSRIEAVKLOMERPMOSKTKIYRPLDRPAS
RSDVNDIKGTPEEISNAOPLRPNAGSS"

BASE COUNT 842 a 800 c 571 g 717 t
ORIGIN

Query Match 100.0%; Score 201; DB 9; Length 2930;
Best Local Similarity 100.0%; Pred. No. 1.8e-48;
Matches 201; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCTAGCAGCAGTAGTCTCTCAAAATCGAAGAGCTTTAGACTTCTAACCGCTGAAGAGG 60
|||||
Db 2051 CCTAGCAGCAGTAGTCTCTCAAAATCGAAGAGCTTTAGACTTCTAACCGCTGAAGAGG 2110

QY 61 GGAACCTGTTATTTTATAGGGGAAGATGCTGTATTATGTTATATCCGAATCGT 120
|||||

Db 2111 GGGACCTGTTATTTTAAAGGAGAAATGCTGTTATTAATCAATCCGGAATCGT 2170
QY 121 CACTGAGAAAGTTAAAGAAATTCAGATCGAATACAGTACAGAGAGAGCTTCGAAA 180
|||||
Db 2171 CACTGAGAAAGTTAAAGAAATTCAGATCGAATACAGTACAGAGAGAGCTTCGAAA 2230
QY 181 CACTGAGACCTGGGGCCTCCT 201
|||||
Db 2231 CACTGAGACCTGGGGCCTCCT 2251

RESULT 8
AR177269 2946 bp DNA linear PAT 17-DEC-2001
LOCUS Sequence 3 from patent US 6312921.
DEFINITION AR177269
ACCESSION AR177269
VERSION AR177269.1 GI:17919624
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 2946)
AUTHORS Jacobus,K., McCoy,J.M., LaVallie,E.R., Racie,L.A., Evans,C.,
Merberg,D., M.S. and Treacy,M.
TITLE Secreted proteins and polynucleotides encoding them.
JOURNAL Patent: US 6312921-A 3 06-NOV-2001;
FEATURES
source 1..2946
location/Qualifiers
BASE COUNT 858 a 801 c 570 g 717 t
ORIGIN

Query Match 100.0%; Score 201; DB 6; Length 2946;
Best Local Similarity 100.0%; Pred. No.1.8e-48;
Matches 201; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCTAGCAGCAGTAGCTCCTCAAAATCGAAGAGCTTTAGACTTGCTAACCGCTGAAGAGG 60
|||||
Db 2049 CCTAGCAGCAGTAGCTCCTCAAAATCGAAGAGCTTTAGACTTGCTAACCGCTGAAGAGG 2108
|||||
QY 61 GGGAACTGTTATTTTAAAGGAGAAATGCTGTTATTAATCAATCCGGAATCGT 120
|||||
Db 2109 GGGAACTGTTATTTTAAAGGAGAAATGCTGTTATTAATCAATCCGGAATCGT 2168
|||||
QY 121 CACTGAGAAAGTTAAAGAAATTCAGATCGAATACAGTACAGAGAGAGCTTCGAAA 180
|||||
Db 2169 CACTGAGAAAGTTAAAGAAATTCAGATCGAATACAGTACAGAGAGAGCTTCGAAA 2228
|||||
QY 181 CACTGAGACCTGGGGCCTCCT 201
|||||
Db 2229 CACTGAGACCTGGGGCCTCCT 2249

RESULT 9
AX007980 10499 bp DNA linear PAT 06-SEP-2000
LOCUS Sequence 3 from Patent W09967395.
DEFINITION AX007980
ACCESSION AX007980
VERSION AX007980.1 GI:9995677
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 10499)
AUTHORS Perin,J.P., Rieger,F. and Alliel,P.M.
TITLE Nucleic sequence and deduced protein sequence family with human
endogenous retroviral motifs, and their uses
JOURNAL Patent: WO 9967395-A 3 29-DEC-1999;
INST NAT SANTE RECH MED (FR); PERIN JEAN PIERRE (FR); RIEGER
FRANCOIS (FR); ALLIEL PATRICK M (FR)
FEATURES
source 1..10499
location/Qualifiers

/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 3048 a 2676 c 2280 g 2495 t
ORIGIN

Query Match 100.0%; Score 201; DB 6; Length 10499;
Best Local Similarity 100.0%; Pred. No.1.9e-48;
Matches 201; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCTAGCAGCAGTAGCTCCTCAAAATCGAAGAGCTTTAGACTTGCTAACCGCTGAAGAGG 60
|||||
Db 9000 CCTAGCAGCAGTAGCTCCTCAAAATCGAAGAGCTTTAGACTTGCTAACCGCTGAAGAGG 9059
|||||
QY 61 GGGAACTGTTATTTTAAAGGAGAAATGCTGTTATTAATCAATCCGGAATCGT 120
|||||
Db 9060 GGGAACTGTTATTTTAAAGGAGAAATGCTGTTATTAATCAATCCGGAATCGT 9119
|||||
QY 121 CACTGAGAAAGTTAAAGAAATTCAGATCGAATACAGTACAGAGAGAGCTTCGAAA 180
|||||
Db 9120 CACTGAGAAAGTTAAAGAAATTCAGATCGAATACAGTACAGAGAGAGCTTCGAAA 9179
|||||
QY 181 CACTGAGACCTGGGGCCTCCT 201
|||||
Db 9180 CACTGAGACCTGGGGCCTCCT 9200

RESULT 10
AX329572 56093 bp DNA linear PAT 09-JAN-2002
LOCUS Sequence 81 from Patent W00194629.
DEFINITION AX329572
ACCESSION AX329572
VERSION AX329572.1 GI:18102550
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1
AUTHORS Young,P.E., Augustus,M., Carter,K.C., Ebner,R., Endress,G.,
Horligan,S., Soppet,D.R. and Weaver,Z.
TITLE Cancer gene determination and therapeutic screening using signature
JOURNAL Patent: WO 0194629-A 81 13-DEC-2001;
Avalon Pharmaceuticals (US)
FEATURES
source 1..56093
location/Qualifiers
BASE COUNT 16164 a 12346 c 10702 g 16881 t
ORIGIN

Query Match 100.0%; Score 201; DB 6; Length 56093;
Best Local Similarity 100.0%; Pred. No.2.1e-48;
Matches 201; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCTAGCAGCAGTAGCTCCTCAAAATCGAAGAGCTTTAGACTTGCTAACCGCTGAAGAGG 60
|||||
Db 37000 CCTAGCAGCAGTAGCTCCTCAAAATCGAAGAGCTTTAGACTTGCTAACCGCTGAAGAGG 37059
|||||
QY 61 GGGAACTGTTATTTTAAAGGAGAAATGCTGTTATTAATCAATCCGGAATCGT 120
|||||
Db 37060 GGGAACTGTTATTTTAAAGGAGAAATGCTGTTATTAATCAATCCGGAATCGT 37119
|||||
QY 121 CACTGAGAAAGTTAAAGAAATTCAGATCGAATACAGTACAGAGAGAGCTTCGAAA 180
|||||
Db 37120 CACTGAGAAAGTTAAAGAAATTCAGATCGAATACAGTACAGAGAGAGCTTCGAAA 37179
|||||
QY 181 CACTGAGACCTGGGGCCTCCT 201
|||||
Db 37180 CACTGAGACCTGGGGCCTCCT 37200

RESULT 11

HSAC000064 56093 bp DNA linear PRI 13-NOV-1996
 LOCUS HSAC000064
 DEFINITION Human BAC clone RG083M05 from 7q21-7q22, complete sequence.
 ACCESSION AC000064
 VERSION AC000064.1 GI:1669369
 KEYWORDS HTG.
 SOURCE Homo sapiens.
 ORGANISM Homo sapiens.
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 REFERENCE 1 (bases 1 to 56093)
 PAULEY, A.
 TITLE The sequence of H. sapiens BAC clone RG083M05
 JOURNAL Unpublished (1996)
 REFERENCE 2 (bases 1 to 56093)
 WATERSTON, R.
 TITLE Direct Submission
 JOURNAL Submitted (13-NOV-1996)
 COMMENT Genome Sequencing Center
 Department of Genetics, Washington University
 St. Louis, MO 63108, USA
 e-mail: saplens@watson.wustl.edu

NOTICE: This sequence may not represent the entire insert of this
 clone. It may be shorter because we only sequence overlapping
 sections once, or longer because we provide a small overlap between
 neighboring submissions.

This sequence was finished as follows unless otherwise noted:
 all regions were double stranded or sequenced with an alternate
 chemistry; an attempt was made to resolve all sequencing problems,
 such as compressions and repeats; all regions were covered by
 sequence from more than one subclone; and the assembly was
 confirmed by restriction digest.

SOURCE INFORMATION:

This clone is from the first release of the human BAC library. The
 library contains cloned DNA from a human male fibroblast cell line
 978SK. For references see: Shizuya et al., Proc. Natl. Acad. Sci.
 89:8794-8797 (1992); Kim et al., Genomics 34:213-218 (1996).
 VECTOR: pBELO
 Selection: chloramphenicol

NEIGHBORING SEQUENCE INFORMATION:

The orientation of this clone is unknown. Actual start of this
 clone is at base position 1 of H_RG083M05; actual end is at 56093
 of H_RG083M05

This clone contains STS SW631725.

FEATURES

Location/Qualifiers

1..56093
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="7"
 /map="7q21-7q22"
 /clone="H_RG083M05"
 /clone_lib="CITB-978SK-B"
 complement(838..1131)
 /rpt_family="ALU"
 <1360..16971
 /gene="WUGSC:H_RG083M05.1"
 join(<1360..1503,4181..4370,4587..4774,6422..6556,
 9483..9547,11631..11773,11864..12021,13131..13286,
 14885..14988,16349..16546,16837..16971)
 /gene="WUGSC:H_RG083M05.1"
 /note="ATPase; strong similarity to peroxisome
 biosynthesis protein PAB1 (PID:91172019); coded for by
 human cDNA C04279 (NID:91467530)"
 /protein_id="AAB46346.1"
 /db_xref="GI:1669371"
 /translation="KRLNIQKTELEVAESEAWMOPSVLLDDLDLGLPAVPEHEH
 SPDAERCEILCNVINKLDCDINKFTDLDLQHVAKETGEGFVARDFTVLVRAHSRL

repeat_region complement(4948..5130)
 /rpt_family="ALU"
 repeat_region complement(6581..7133)
 /rpt_family="L1"
 repeat_region complement(7767..8037)
 /rpt_family="ALU"
 repeat_region complement(8186..8472)
 /rpt_family="ALU"
 misc_feature 8473..8625
 /gene="WUGSC:H_RG083M05.1"
 /note="match to human 3' EST H75782 (NID:91049794), bases
 287-444"
 8841..9161
 /gene="WUGSC:H_RG083M05.1"
 /note="match to human 5' EST H75921 (NID:91050050), bases
 21-348"
 9481..9547
 /gene="WUGSC:H_RG083M05.1"
 /note="match to human 5' EST N22627 (NID:91130501), bases
 276-343"
 complement(12612..12907)
 /rpt_family="ALU"
 misc_feature 13670..13793
 /gene="WUGSC:H_RG083M05.1"
 /note="match to human 5' EST H41382 (NID:917434), bases
 143-266"
 13794..13877
 /rpt_family="ALU"
 misc_feature 13878..13906
 /gene="WUGSC:H_RG083M05.1"
 /note="match to human 5' EST H41382 (NID:917434), bases
 30-58"
 13907..14104
 /rpt_family="ALU"
 repeat_region complement(14110..14137)
 /rpt_family="L1"
 repeat_region complement(15618..15907)
 /rpt_family="ALU"
 repeat_region 17227..17522
 /rpt_family="ALU"
 misc_feature 18667..19235
 /note="match to human fetal brain 5' EST D61494
 (NID:970409), bases 1-255, and to human 3' EST R07476
 (NID:9759399)"
 19550..19670
 /rpt_family="ALU"
 misc_feature 21507..37303
 /note="similarity to various ss-RNA virus polypeptides;
 pseudogene; region of matches and close matches to
 multiple human ESTs, see R66740 (NID:9842257)"
 37316..37489
 /note="Grail prediction, score = 80"
 /evidence="not_experimental"
 repeat_region complement(38938..39224)
 /rpt_family="ALU"
 misc_feature 39225..39707
 /note="match to multiple human ESTs, see N30113
 (NID:91148633)"
 39800..40085
 /rpt_family="ALU"
 repeat_region complement(40247..40538)
 /rpt_family="ALU"
 repeat_region complement(40632..40924)
 /rpt_family="ALU"
 repeat_region complement(42283..42891)

```

repeat_region /rpl_family="ALU"
complement(45474..45613)
misc_feature /rpl_family="ALU"
complement(45614..45737)
/ note="match to human 3' EST H48898 (NID:g988738), bases
129-333"
misc_feature complement(46107..47026)
/ note="match to multiple human ESTs, see N81064
(NID:g1243765), H48897 (NID:g988737), and M78831
(NID:g273146)"
repeat_region complement(47027..47318)
/ rpl_family="ALU"
misc_feature complement(47365..47782)
/ note="match to multiple human ESTs, see W37495
(NID:g1319089)"
misc_feature 47898..48115
/ note="match to human 5' EST H62306 (NID:g1015138), bases
93-368"
repeat_region complement(48116..48405)
/ rpl_family="ALU"
misc_feature complement(48406..48584)
/ note="match to human 3' EST N29952 (NID:g1148472), bases
290-455, and 5' EST R12730 (NID:g765806)"
repeat_region complement(48787..49405)
/ rpl_family="ALU"
misc_feature complement(49406..49534)
/ note="match to human 3' EST R65794 (NID:g838432), bases
309-440"
repeat_region complement(49638..49672)
/ rpl_family="ALU"
misc_feature complement(49674..49890)
/ note="match to human 3' EST N29952 (NID:g1148472) and 5'
EST N29938 (NID:g1148458), sequences are from opposite
ends of the same clone"
gene complement(49698..51806)
/ gene="WUGSC:H.RG083M05.2"
complement(join(49698..49888,51575..51806))
/ gene="WUGSC:H.RG083M05.2"
/ note="coded for by human CDNA5 W37389 (NID:g1319205),
R65891 (NID:g838529), R65794 (NID:g838432) and R65794
(NID:g838432)"
/ codon_start=1
/ protein_id="AAB6345.1"
/ db_xref="GI:1669370"
/ translation="MEFYRPGGIIICFPGVYVYQGVYSDVDEQGRYYAQIRF
IQDYCEKSAALWILPTLSSPRDQEDPASTIGPEEDLFRKMEYLEFVCHAPSEIFK
SRSSPPVPTREKGYIWHVGPPTPATIKESVANHL"
exon complement(51576..51758)
/ gene="WUGSC:H.RG083M05.2"
/ note="Grail prediction, score = 86"
/ evidence=not-experimental
repeat_region complement(52052..52329)
/ rpl_family="L1"
misc_feature 55557..55843
/ note="match to human EST M79192 (NID:g273505) base 2-289"

```

```

Db 37180 CACTGAGACCTGGGGCCTCCT 37200
|||||
RESULT 12
AC007566/c 149194 bp DNA linear PRI 01-MAR-2002
LOCUS Homo sapiens BAC clone CTB-10G5 from 7q21-7q22, complete sequence.
DEFINITION AC007566
ACCESSION AC007566
VERSION AC007566.2 GI:11181861
KEYWORDS HTG.
SOURCE Homo sapiens.
ORGANISM Homo sapiens.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
1 (bases 1 to 149194)
Sulston, J.E. and Waterston, R.
Toward a complete human genome sequence
Genome Res. 8 (11), 1097-1108 (1998)
99063792
PUBMED 9847074
2 (bases 1 to 149194)
Du, Z.
The sequence of Homo sapiens BAC clone CTB-10G5
Unpublished (2001)
3 (bases 1 to 149194)
Waterston, R.H.
Direct Submission
Submitted (15-MAY-1999) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
4 (bases 1 to 149194)
Waterston, R.
Direct Submission
Submitted (02-OCT-2000) Department of Genetics, Washington
University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
5 (bases 1 to 149194)
Waterston, R.H.
Direct Submission
Submitted (16-NOV-2000) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
6 (bases 1 to 149194)
Waterston, R.H.
Direct Submission
Submitted (03-JUN-2002) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
7 (bases 1 to 149194)
Waterston, R.H.
Direct Submission
Submitted (06-FEB-2002) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
8 (bases 1 to 149194)
Waterston, R.
Direct Submission
Submitted (01-MAR-2002) Department of Genetics, Washington
University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
On Nov 16, 2000 this sequence version replaced gi:4835815.
----- Genome Center
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: http://genome.wustl.edu/gsc
Contact: saplens@wustl.wustl.edu
----- Summary Statistics
Center project name: H_RG010G05

```

NOTICE: This sequence may not represent the entire insert of this clone. It may be shorter because we only sequence overlapping clone sections once, or longer because we provide a small overlap between neighboring data submissions.

QY 121 CACTGAGAAAGTTAAAGAAATTCGATCGAATACAGTAGACAGAGAGCTTCGAAA 180
|||||
Db 84884 CACTGAGAAAGTTAAAGAAATTCGATCGAATACAGTAGACAGAGAGCTTCGAAA 84825
QY 181 CACTGACCCCTGGGGCCTCCT 201
|||||
Db 84824 CACTGACCCCTGGGGCCTCCT 84804

RESULT 13
AF156963 1617 bp DNA linear PRI 06-JUN-2000
AF156963
LOCUS Homo sapiens human endogenous retrovirus W envC7-1 envelope protein
DEFINITION (env) gene, complete cds.
ACCESSION AF156963
VERSION AF156963.1 GI:8272467
KEYWORDS
SOURCE Homo sapiens.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 1617)
AUTHORS Voisset,C., Bouton,O., Bedin,F., Duret,L., Mandrand,B., Mallet,F.
and Paranhos-Baccala,G.
TITLE Chromosomal distribution and coding capacity of the human
JOURNAL endogenous retrovirus HERV-W family
MEDLINE AIDS Res. Hum. Retroviruses 16 (8), 731-740 (2000)
PUBMED 20284713
10826480
REFERENCE 2 (bases 1 to 1617)
AUTHORS Voisset,C., Bouton,O., Bedin,F., Duret,L., Mandrand,B., Mallet,F.
and Paranhos-Baccala,G.
TITLE Direct Submission
JOURNAL Submitted (09-JUN-1999) UMR103CNRS-biomerieux, ENS Lyon, 46 allée
d'Italie, Lyon 69364, France
FEATURES
source
1. 1617
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="7"
/clone="envC7-1"
/note="human/rodent monochromosomal NIGMS somatic hybrid
panel #2 (Coriell Institute, Camden, NJ, USA); clone
GM/NA10791 retaining human chromosome 7"
1. 1617
/gene="env"
CDS
1. 1617
/gene="env"
/codon_start=1
/product="envelope protein"
/protein_id="AA074215.1"
/db_xref="GI:8272468"
/translation="MALPHTLFTVLLPSFTLTAPPSCRMSSSPYQEFLMRQRP
GNIDAPYRSLSKGTPTPTATHTPRNCYSATLCMAHNTHTGKMINPSCGLGV
TVCWTFYQTGMSDGGVODAREKHVEYSOLTRVHGTSPPKGLDSLHETLRT
HTRLVSLPNTTLTGLEHVSANPNCWICPLNRPVSIIPVEQNNPSEIHTTSV
LVGPVLSNLETHTSNLTGVKESNTYTTSNOCIRWTPPTQIVCSGIFVCGTSA
YKRLNSESSEKCLSPFLVPPTITTEDLYNYSKPRNKRPVLLPVTIGAGVGAAG
TGIGGTTSTQFYKLSQELNGDMERADSLVTLQDOLNSIAAVLONRRALDLITAE
RGCTCLFLGECYVYVNOGIVTEKVEIRDIORRAEBELNTPGGLSOMMWTIP
FLGPALAIILLFLGPCIFNLIVFVFSRIEAVKLQMEPKMQSTKTYRRPLDRPASP
RSDVNDIKGTPPEISAAQPLRRNSAGSS"
repeat_region
1. 1617
/rpt_family="human endogenous retrovirus W"
/rpt_type="dispersed"
1. 18
/gene="env"
primer_bind
1597. 1617
/gene="env"
BASE COUNT 441 a 452 c 297 g 427 t
ORIGIN
Query Match 99.2%; Score 199.4; DB 9; Length 1617;

Best Local Similarity 99.5%; Pred. No. 5,1e-48;
Matches 200; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 CCTAGCAGCAGTAGTCTCTTAAATCGAAGAGCTTTAGACTTGCTAACCGCTGAAGAG 60
|||||
Db 1122 CCTAGCAGCAGTAGTCTCTTAAATCGAAGAGCTTTAGACTTGCTAACCGCTGAAGAG 1181
QY 61 GGGAACTGTTATTTTAGGGGAAGATGCTGTTATATGTTATGCAATCGGAATCGT 120
|||||
Db 1182 GGGAACTGTTATTTTAGGGGAAGATGCTGTTATATGTTATGCAATCGGAATCGT 1241
QY 121 CACTGAGAAAGTTAAAGAAATTCGATCGAATACAGTAGACAGAGAGCTTCGAAA 180
|||||
Db 1242 CACTGAGAAAGTTAAAGAAATTCGATCGAATACAGTAGACAGAGAGCTTCGAAA 1301
QY 181 CACTGACCCCTGGGGCCTCCT 201
|||||
Db 1302 CACTGACCCCTGGGGCCTCCT 1322

RESULT 14
AF506835 1624 bp mRNA linear PRI 20-MAY-2002
AF506835
LOCUS Homo sapiens envverin mRNA, partial cds.
DEFINITION AF506835
ACCESSION AF506835
VERSION AF506835.1 GI:20978307
KEYWORDS
SOURCE Homo sapiens.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 1624)
AUTHORS Alliel,P.M., Perin,J.P., Pierly,R., Nusbaum,J.L., Menard,A. and
Rieger,F.
TITLE Endogenous retroviruses and multiple sclerosis. II. HERV-7q
JOURNAL C. R. Acad. Sci. III, Sci. Vie 321 (10), 857-863 (1998)
MEDLINE 99052087
PUBMED 98350022
REFERENCE 2 (bases 1 to 1624)
AUTHORS Alliel,P.M., Perin,J.P., Goudou,D., Bitoun,M., Robert,B. and
Rieger,F.
TITLE The HERV-W/q family in the human genome. Potential for protein
JOURNAL expression and gene regulation
REFERENCE 3 (bases 1 to 1624)
AUTHORS Goudou,D., Perin,J.P., Rieger,F., Robert,B. and Alliel,P.M.
TITLE Direct Submission
JOURNAL Submitted (29-APR-2002) U-488, INSERM, 80, rue du General Leclerc,
Le Kremlin-Bicetre 94270, France
FEATURES
source
1. 1624
/organism="Homo sapiens"
/db_xref="taxon:9606"
/db_xref="GI:20978308"
/tissue_type="placenta"
25. >1624
/codon_start=1
/product="envverin"
/protein_id="AA033413.1"
/db_xref="GI:20978308"
/translation="MALPHTLFTVLLPSFTLTAPPSCRMSSSPYQEFLMRQRP
GNIDAPYRSLSKGTPTPTATHTPRNCYSATLCMAHNTHTGKMINPSCGLGV
TVCWTFYQTGMSDGGVODAREKHVEYSOLTRVHGTSPPKGLDSLHETLRT
HTRLVSLPNTTLTGLEHVSANPNCWICPLNRPVSIIPVEQNNPSEIHTTSV
LVGPVLSNLETHTSNLTGVKESNTYTTSNOCIRWTPPTQIVCSGIFVCGTSA
YKRLNSESSEKCLSPFLVPPTITTEDLYNYSKPRNKRPVLLPVTIGAGVGAAG
TGIGGTTSTQFYKLSQELNGDMERADSLVTLQDOLNSIAAVLONRRALDLITAE
RGCTCLFLGECYVYVNOGIVTEKVEIRDIORRAEBELNTPGGLSOMMWTIP
FLGPALAIILLFLGPCIFNLIVFVFSRIEAVKLQMEPKMQSTKTYRRPLDRPASP
RSDVNDIKGTPPEISAAQPLRRN"
BASE COUNT 449 a 456 c 290 g 429 t
ORIGIN
Query Match 99.2%; Score 199.4; DB 9; Length 1624;

Best Local Similarity 99.5%; Pred. No. 5.1e-48;
Matches 200; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CCTAGCAGCAGTAGTCCTTCAAAATCGAAGAGCTTTAGACTTGCTAACCGCTGAAGAGG 60
|||||
Db 1146 CTTAGCAGCAGTAGTCCTTCAAAATCGAAGAGCTTTAGACTTGCTAACCGCTGAAGAGG 1205

QY 61 GGGACCTGTTTATTTTATTTAGGGAGAGATCTGTTATTTATTTATCAATCCGGAATCGT 120
|||||
Db 1206 GGGACCTGTTTATTTTATTTAGGGAGAGATCTGTTATTTATTTATCAATCCGGAATCGT 1265

QY 121 CACTGAGAAAGTTAAAGAAATTCGAGATCGAATPACAACGTAAGCAGAGAGAGCTTCGAAA 180
|||||
Db 1266 CACTGAGAAAGTTAAAGAAATTCGAGATCGAATPACAACGTAAGCAGAGAGAGCTTCGAAA 1325

QY 181 CACTGACCTCGGGGCTCCT 201
|||||
Db 1326 CACTGACCTCGGGGCTCCT 1346

RESULT 15
AF513360 1860 bp mRNA linear PRI 05-JUN-2002
LOCUS Homo sapiens enverin mRNA, complete cds.
DEFINITION AF513360
ACCESSION AF513360.1 GI:21326140
VERSION
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 1860)
AUTHORS Alilel, P.M., Perin, J.P., Pierig, R., Nusbaum, J.L., Menard, A. and Rieger, F.
TITLE Endogenous retroviruses and multiple sclerosis. Part 2: HERV-7g and its env transcripts
JOURNAL C. R. Acad. Sci. III, Sci. Vie 312, 857-863 (1998)
REFERENCE 2 (bases 1 to 1860)
AUTHORS Alilel, P.M., Perin, J.P., Goudou, D., Bitoun, M., Robert, B. and Rieger, F.
TITLE The HERV-W/7g family in the human genome. Potential for protein expression and gene regulation
JOURNAL Cell Mol. Biol. 48 (2), 213-217 (2002)
MEDLINE 21985840
PUBMED 11990458

REFERENCE 3 (bases 1 to 1860)
AUTHORS Alilel, P.M., Goudou, D., Perin, J.P. and Rieger, F.
TITLE Direct Submission
JOURNAL Submitted (18-MAY-2002) U-488, INSERM, 80, rue du General Leclerc,
Le Kremlin-Bicetre 94270, France

FEATURES
source
1. 1860
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/map="7q21"
/issue_type="placenta"
59.1675
/codon_start=1
/product="enverin"
/protein_id="AA047599.1"
/db_xref="GI:21326141"
/translation="MALPYHILFTVLPSPFTLAPPCRCMTSSSPYQFLMRMRP
GNIDAPYSRLSKGPTPTAHTHMPKNCYHSATLCMHANTHYWTGKMINPSCDGLGV
TYCWTFYQOTGSDGGVODDA REKHUKVVISQLTRVHTSSPYGLDLSKHEITLT
HTRIVSLENTITLTGLHEVSAQNPNCWITCLPLNFRYVISIPVEQMNNSSTEINTSV
LVGPLVSNLEITHNSLVCVKSNTYTTNSOCIRWPPPTQVCLPSGIFVCGTSA
VYCLNGSSSMCFSLFLVPPMTIYTEDOLYNYVSKPRNKRVPDILPFVIGAGVLGATG
TGIGITTSYFYKLSQELNDGMDERVADSLVTLQDLSLAAYVIONRRLADLLTAE
RGGCTTLFGECCYVYVNOGSIYTERKVEIRDRIORAEELRNTGPMGLISOMPWLP
FLGPLAAILILLRGCTCFNLIVNFVSSRIEAVKIQMEPKMQSKTKIYRPLDRASP
RSDVDIKGTPEEISAOPFLRPSAGSS"

CDS

BASE COUNT 513 a 520 c 350 g 477 t

ORIGIN
Query Match 99.2%; Score 199.4; DB 9; Length 1860;
Best Local Similarity 99.5%; Pred. No. 5.1e-48;
Matches 200; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CCTAGCAGCAGTAGTCCTTCAAAATCGAAGAGCTTTAGACTTGCTAACCGCTGAAGAGG 60
|||||
Db 1180 CTTAGCAGCAGTAGTCCTTCAAAATCGAAGAGCTTTAGACTTGCTAACCGCTGAAGAGG 1239

QY 61 GGGACCTGTTTATTTTATTTAGGGAGAGATCTGTTATTTATTTATCAATCCGGAATCGT 120
|||||
Db 1240 GGGACCTGTTTATTTTATTTAGGGAGAGATCTGTTATTTATTTATCAATCCGGAATCGT 1299

QY 121 CACTGAGAAAGTTAAAGAAATTCGAGATCGAATPACAACGTAAGCAGAGAGAGCTTCGAAA 180
|||||
Db 1300 CACTGAGAAAGTTAAAGAAATTCGAGATCGAATPACAACGTAAGCAGAGAGAGCTTCGAAA 1359

QY 181 CACTGACCTCGGGGCTCCT 201
|||||
Db 1360 CACTGACCTCGGGGCTCCT 1380

Search completed: May 2, 2003, 15:27:56
Job time : 756 secs

GenCore version 5.1.5
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 2, 2003, 12:49:35 ; Search time 159.333 Seconds

(without alignments)
2840.910 Million cell updates/sec

Title: US-09-719-554-3_COPY_9000_9200

Perfect score: 201
Sequence: 1 cctagcagcagctagctctc.....actgacacctgagcctctct 201

Scoring table: IDENTITY-NUC
Gapop 10.0 , Gapept 1.0

Searched: 2185239 seqs, 112599159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :
1: N.Geneseq_101002: *
2: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT: *
3: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT: *
4: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT: *
5: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT: *
6: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT: *
7: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT: *
8: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT: *
9: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT: *
10: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT: *
11: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT: *
12: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT: *
13: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT: *
14: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT: *
15: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT: *
16: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT: *
17: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT: *
18: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT: *
19: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT: *
20: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT: *
21: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT: *
22: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT: *
23: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT: *
24: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	201	100.0	1617	22	AAH20070
2	201	100.0	2055	21	ABN97948
3	201	100.0	2599	21	ABN97927
4	201	100.0	2781	22	AA555530
5	201	100.0	2930	24	AA24195
6	201	100.0	2946	20	AA77526
7	201	100.0	2946	21	AA59468
8	201	100.0	10499	21	ABN97929
9	201	100.0	56093	24	AB161744

10	198.2	98.6	7582	20	AA525665	Complete human end
11	198.2	98.6	7582	21	AA525665	Human endogenous r
12	197.8	98.4	1136	20	AA525660	Human endogenous r
13	197.8	98.4	1136	21	AA525660	3' pol gene and 3'
14	197.8	98.4	6394	23	AA584210	DNA encoding novel
15	193	96.0	2782	20	AA525661	Human endogenous r
16	193	96.0	2782	21	AA525661	5' non coding, 3'
17	193	96.0	2782	22	AAH20069	HERV-W envelope pr
18	183.4	91.2	591	20	AA525685	Human endogenous r
19	183.4	91.2	591	21	AA525685	Probe PenV-C15 use
20	183.4	91.2	1481	19	AA543217	Multiple sclerosis
21	183.4	91.2	1481	20	AA529703	Clone C15 from MSR
22	183.4	91.2	1629	21	AA566625	DNA encoding an en
23	183.4	91.2	2030	21	AA563826	Nucleotide sequenc
24	180.4	89.8	1088	23	AA592689	DNA encoding novel
25	180.4	89.8	1174	23	AA592498	DNA encoding novel
26	180.4	89.8	1929	23	AA571718	DNA encoding novel
27	180.4	89.8	1929	23	AA576466	DNA encoding novel
28	180.4	89.8	1929	23	AA592493	DNA encoding novel
29	180.4	89.8	1929	23	AA592666	DNA encoding novel
30	178.6	88.9	1894	22	ABA43822	Human breast cell
31	178.6	88.9	1894	22	ABA56337	Human foetal liver
32	178.6	88.9	1894	22	ABA25978	Probe #4444 for ge
33	178.6	88.9	1894	22	AAK04516	Human brain expres
34	178.6	88.9	1894	22	AAK30018	Human bone marrow
35	178.6	88.9	1894	22	AA114608	Probe #4541 for ge
36	178.6	88.9	1894	22	AA135960	Probe #4666 used t
37	178.6	88.9	1894	22	AA104422	Probe #4413 used t
38	178.6	88.9	1894	24	AB504589	Human genome-deriv
39	178.6	88.9	2784	21	ABN97930	Human retroviral s
40	164.2	81.7	1446	23	AA567605	DNA encoding novel
41	164.2	81.7	1446	23	AA573545	DNA encoding novel
42	164.2	81.7	1721	23	AA572233	DNA encoding novel
43	163.8	81.5	1329	19	AA543219	Multiple sclerosis
44	163.8	81.5	1329	20	AA529704	Clone 5M6 from MSR
45	156.2	77.7	902	23	AA571722	DNA encoding novel

ALIGNMENTS

RESULT 1	AAH20070	standard; DNA: 1617 BP.
ID	AAH20070	
AC	AAH20070:	
XX	XX	
DE	08-AUG-2001	(first entry)
XX	XX	
DE	HERV-W envelope protein G coding sequence.	
XX	XX	
KW	Human endogenous retrovirus; HERV-W; chromosome 7; env protein;	
KW	envelope protein; multiple sclerosis-related superantigen; vaccine;	
KW	surface antigen; transmembrane; multiple sclerosis; neuroprotective;	
KW	antisense-therapy; autoimmune disorder; ds.	
XX	XX	
OS	Human endogenous retrovirus.	
XX	XX	
FT	Key	Location/Qualifiers
FT	CDS	1..1617
FT	FT	/product= "HERV-W envelope protein G"
XX	XX	
PN	W0200131021-A1.	
XX	XX	
PD	03-MAY-2001.	
XX	XX	
PF	30-OCT-2000; 2000WO-EP10659.	
XX	XX	
PR	28-OCT-1999; 99EP-0402690.	
XX	XX	
PA	(UYGE-) UNIV GENEVE.	
XX	XX	

PI Conrad B, Mach B;
XX WPI; 2001-316336/33.
DR P-PSDB; AAB75138.
XX
PT New human retrovirus HERV-W ENV proteins/peptides having superantigen
XX activity useful for diagnosing and treating multiple sclerosis -
XX
PS Claim 13; Fig 10; 94pp; English.
XX
CC On the basis of the PBS t-RNA motif used for the classification of human
CC endogenous retrovirus (HERVs) the full length endogenous provirus which
CC was located on the long arm of human chromosome 7 (7q21.22) has been
CC designated HERV-W. The present invention describes proteins or peptides
CC (I) having superantigen (Sag) activity comprising the ENV protein (ENV)
CC of HERV-W, the surface protein (SU) and transmembrane (TM) sub-units. (I)
CC have neuroprotective activity, and can be used in: vaccines; antisense
CC therapy; and HERV-W Sag activity-inhibitors. (I) and encoding DNA/RNA are
CC useful for diagnosing multiple sclerosis (MS) or HERV-W-associated
CC disorders. (I) are also useful for identifying substances (and optionally
CC recovering) capable of binding to a retroviral superantigen associated
CC with MS, substances capable of blocking Sag activity and substances
CC capable of blocking transcription or translation of HERV-W retroviral
CC superantigen. A protein or peptide derived from (I), modified to be
CC devoid of Sag activity and being capable of generating an immune response
CC against HERV-W retroviral Sag is useful in therapy. Nucleic acid
CC molecules encoding (I) are useful as vaccines against MS. Substances
CC capable of blocking Sag activity, capable of binding to a retroviral
CC superantigen associated with MS, or capable of blocking transcription or
CC translation of HERV-W retroviral superantigen for use in treating or
CC preventing MS, obtained using (I) are useful for the treatment and
CC prevention of MS. (I) and nucleic acids encoding them are useful for
CC diagnosing autoimmune disease. The present sequence encodes the
CC specifically claimed envelope protein of HERV-W designated G.
XX
SQ Sequence 1617 BP; 442 A; 452 C; 296 G; 427 T; 0 other;
XX
Query Match 100.0%; Score 201; DB 22; Length 1617;
Best Local Similarity 100.0%; Pred. No. 4.9e-55;
Matches 201; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
QY 1 CCGAGACGAGTGTCTTCAAAATGCAAGAGCTTAGACTCTTAACCGCTGAAGAGG 60
DB 1122 CCGAGCAGCAGTAGCTCTTCAAAATGCAAGAGCTTAGACTCTTAACCGCTGAAGAGG 1181
QY 61 GGGACCTGTTATTTTAAAGGAGAGATGCTGTATTATGTTAATCAATCCGGAATCGT 120
DB 1182 GGGACCTGTTATTTTAAAGGAGAGATGCTGTATTATGTTAATCAATCCGGAATCGT 1241
QY 121 CACTGAGAAAGTTAAAGAAATTCGAGATCGAATACAGTAGAGAGAGAGCTTGAAA 180
DB 1242 CACTGAGAAAGTTAAAGAAATTCGAGATCGAATACAGTAGAGAGAGAGCTTGAAA 1301
QY 181 CACTGACCTGGGGGCTCCT 201
DB 1302 CACTGACCTGGGGGCTCCT 1322

RESULT 2
ABN97948
ID ABN97948 standard; DNA; 2055 BP.
XX
AC ABN97948;
XX
DT 01-AUG-2002 (first entry)
XX
DE Human retroviral HERV-7q env coding sequence.
XX
XX Autoimmune disease; HERV-7q; chromosome 7q; immunotherapy;
KW multiple sclerosis; ds.
XX
OS Human retrovirus.
XX

PN WO9967395-A1.
XX
XX 29-DEC-1999.
PD
XX
PF 23-JUN-1999; 99WO-FR01513.
XX
PR 23-JUN-1998; 98FR-0007920.
XX
PA (INRM) INSERM INST NAT SANTE & RECH MEDICALE.
XX
XX Allel PM, Perin J, Rieger F;
PI
DR WPI; 2000-160587/14.
XX
PT New nucleic acid sequences of human endogenous retrovirus, HERV-7q,
PT used for diagnosis, treatment and prevention of autoimmune and
PT neurological diseases -
XX
PS Claim 3; Page 142-145; 225pp; French.
XX
CC The present invention relates to new nucleic acid sequences of human
CC endogenous retrovirus, HERV-7q, which is located on chromosome 7q.
CC Regulatory elements associated with HERV-7q may alter expression of other
CC genes (even remote genes) on the same chromosome, inducing immunological
CC and/or neurological changes (which may be pathological or protective/
CC curative). HERV-7q peptides can be used to improve efficiency of the
CC immune response, e.g. in immunotherapy. HERV-7q peptides and their coding
CC sequences can be used in immunogenic or vaccinating compositions, for
CC protection against autoimmune diseases, particularly multiple sclerosis.
CC The peptides may also be used (by sequence comparison) to detect/identify
CC endogenous retroviruses that are abnormally expressed in cancer,
CC neuropathologies or other autoimmune diseases. The present sequence was
CC used to illustrate the invention.
XX
SQ Sequence 2055 BP; 576 A; 574 C; 376 G; 529 T; 0 other;
XX
Query Match 100.0%; Score 201; DB 21; Length 2055;
Best Local Similarity 100.0%; Pred. No. 5.4e-55;
Matches 201; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
QY 1 CCGAGACGAGTGTCTTCAAAATGCAAGAGCTTAGACTCTTAACCGCTGAAGAGG 60
DB 1512 CCGAGCAGCAGTAGCTCTTCAAAATGCAAGAGCTTAGACTCTTAACCGCTGAAGAGG 1571
QY 61 GGGACCTGTTATTTTAAAGGAGAGATGCTGTATTATGTTAATCAATCCGGAATCGT 120
DB 1572 GGGACCTGTTATTTTAAAGGAGAGATGCTGTATTATGTTAATCAATCCGGAATCGT 1631
QY 121 CACTGAGAAAGTTAAAGAAATTCGAGATCGAATACAGTAGAGAGAGAGCTTGAAA 180
DB 1632 CACTGAGAAAGTTAAAGAAATTCGAGATCGAATACAGTAGAGAGAGAGCTTGAAA 1691
QY 181 CACTGACCTGGGGGCTCCT 201
DB 1692 CACTGACCTGGGGGCTCCT 1712

RESULT 3
ABN97927
ID ABN97927 standard; DNA; 2599 BP.
XX
AC ABN97927;
XX
DT 01-AUG-2002 (first entry)
XX
DE Human retroviral sequence HERV 7 env.
XX
XX Autoimmune disease; HERV-7q; chromosome 7q; immunotherapy;
KW multiple sclerosis; ds.
XX
OS Human retrovirus.
XX
PN WO9967395-A1.

XX 29-DEC-1999.
PD
PF 23-JUN-1999; 99WO-FR01513.
XX
PR 23-JUN-1998; 98FR-0007920.
XX
PA (INRM) INSEPM INST NAT SANTE & RECH MEDICALE.
XX
PI Alliel PM, Perin J, Rieger F;
XX WPI; 2000-160587/14.
DR
XX New nucleic acid sequences of human endogenous retrovirus, HERV-7q,
PT used for diagnosis, treatment and prevention of autoimmune and
PT neurological diseases
XX
PS Claim 1; Page 128-129; 225pp; French.
XX
CC The present invention relates to new nucleic acid sequences of human
CC endogenous retrovirus, HERV-7q, which is located on chromosome 7q.
CC Regulatory elements associated with HERV-7q may alter expression of other
CC genes (even remote genes) on the same chromosome, inducing immunological
CC and/or neurological changes (which may be pathological or protective/
CC curative). HERV-7q peptides can be used to improve efficiency of the
CC immune response, e.g. in immunotherapy. HERV-7q peptides and their coding
CC sequences can be used in immunogenic or vaccinating compositions, for
CC protection against autoimmune diseases, particularly multiple sclerosis.
CC The peptides may also be used (by sequence comparison) to detect/identify
CC endogenous retroviruses that are abnormally expressed in cancer,
CC neuropathologies or other autoimmune diseases. The present sequence was
CC used to illustrate the invention.
XX
SQ Sequence 2599 BP; 744 A; 718 C; 495 G; 642 T; 0 other;

	Query Match	Similarity	100.0%	Score	201:	DB	21:	Length	2599;	
	Best Local	Similarity	100.0%	Pred.	No. 5.8e-55;					
	Matches	201:	Conservative	0;	Mismatches	0;	Indels	0;	Gaps	0;
Dy	1	CCTAGCAGCAGTAGTCCTCCAAATCGAAGAAGCTTTAGACTGTCAACCGCTGAAGAAGG	60							
Db	2050	CCTAGCAGCAGTAGTCCTCCAAATCGAAGAAGCTTTAGACTGTCAACCGCTGAAGAAGG	2109							
Dy	61	GGGAACCTGTTATTATTTTAAAGGGAAGAAATGCTGTTATTATGTTAATCAAATCCGAATCGT	120							
Db	2110	GGGAACCTGTTATTATTTTAAAGGGAAGAAATGCTGTTATTATGTTAATCAAATCCGAATCGT	2169							
Dy	121	CACGTAGAAAAGTTAAAGAAATTCGAGATCGAATTACAAGTAAAGAGAGAGACTTCGAAA	180							
Db	2170	CACGTAGAAAAGTTAAAGAAATTCGAGATCGAATTACAAGTAAAGAGAGAGACTTCGAAA	2229							
Dy	181	CACGTGACCCTGGGGGCTCTCT	201							
Db	2230	CACGTGACCCTGGGGGCTCTCT	2250							
	RESULT 4									
	AAF55630									
	ID	AAF55630 standard; DNA; 2781 BP.								
	xx	AAF55630;								
	xx									
	xx	29-MAY-2001 (first entry)								
	xx									
	DE	Nucleotide sequence of a human endogenous retrovirus envelope protein.								
	xx									
	KW	Envelope protein; HERV; syncytia formation; placental development;								
	xw	syncytia; cancer; cell adhesion; ss.								
	xx									
	OS	Human endogenous retrovirus.								
	xx									
	Key	Location/Qualifiers								
	FH									
	CDS	762..2378								

FT	/*tag=	a
FT	/product=	"envelope protein"
XX		
FN	MO200116171-A1.	
XX		
PD	08-MAR-2001.	
PF		
XX	01-SEP-2000; 2000WO-FR02429.	
XX		
PR	01-SEP-1999; 99FR-0011141.	
XX	15-SEP-1999; 99FR-0011793.	
PA	(INMR) BIO MERIEUX.	
PA	(INRM) INST NAT SANTE & RECH MEDICALE.	
PI		
XX	Mallet F, Cosslet F, Blond J, Lavillette D, Boulton O, Ruggieri A;	
DR	WPI: 2001-226676/23.	
DR	P-PSDB; AAB67652.	
XX		
PT	Detecting expression of human endogenous retrovirus envelope protein in	
XX	cells of a tissue or culture, from its ability to induce syncytia -	
PS	Disclosure; Page 44-45; 57pp; French.	
XX		
CC	The present sequence encodes a human endogenous retrovirus envelope	
CC	protein. The specification describes a method for detecting expression	
CC	of an envelope protein from a human endogenous retrovirus (HERV), in	
CC	cells, of a tissue or culture. The method comprises detecting syncytia	
CC	formation due to the fusogenic properties of the envelope protein.	

Query Match	100.0%	Score 201	DB 22	Length 2781
Best Local Similarity	100.0%	Pred. No. 6e-55		
Matches 201	Conservative	0	Mismatches	0
			Indels	0
			Gaps	0
Qy 1	CTTACGACGAGTACGCTCTTCAAAATCGAAGACCTTTAGACTTGTCTAACCGCTGAAGAAGG	60		
Db 1883	CCTACGACGAGTACGCTCTTCAAAATCGAAGACCTTTAGACTTGTCTAACCGCTGAAGAAGG	1942		
Qy 61	GGGAACCTGTTATTTTATTTAGGGGAAGATGCTGTTATTTATTTCAATCCGGAATCGT	120		
Db 1943	GGGAACCTGTTATTTTATTTTATTTAGGGGAAGATGCTGTTATTTATTTCAATCCGGAATCGT	2002		
Qy 121	CACGTAGAAGTAAAGAATTCGAGATCGATACACGTAGACGAGAGAGACTTCGANA	180		
Db 2003	CACGTAGAAGTAAAGAATTCGAGATCGATACACGTAGACGAGAGAGACTTCGANA	2062		
Qy 181	CACGTGACCTGCGGCGCTCTCT	201		
Db 2063	CACGTGACCTGCGGCGCTCTCT	2083		
RESULT 5				
AAD24195				
TD	AAD24195 standard; cDNA; 2930 BP.			
XX				
AC	AAD24195;			
XX				
DT	07-MAY-2002 (first entry)			
XX				
DE	Human syncytin cDNA.			
KW	Human; syncytin; preclampsia; gestational trophoblast disorder; choriocarcinoma; hydatiform mole; placental site tumour; abortion;			

KW envelope gene: human endogenous defective retrovirus; HERV-W; ss.
XX Homo sapiens.
XX Location/Qualifiers
FH Key 930..2546
FT CDS /*tag=a
FT /product="Syncytin"
XX
XX MO200204678-A2.
XX
XX 17-JAN-2002.
XX
XX 09-JUL-2001; 2001WO-US21719.
XX
XX 07-JUL-2000; 2000US-216657P.
XX
XX (GENY) GENETICS INST INC.
XX
XX Keith JC, McCoy JM, M1 S;
XX
XX WPI: 2002-171727/22.
XX P-PSDB; AAE14540.
XX
XX Identifying a compound for treating a subject with or at risk of
XX developing preeclampsia, comprises determining whether the expression
XX or activity of syncytin in the cell is modulated in the presence of a
XX test compound
XX
XX
XX Disclosure: Page 39-42; 43pp; English.
XX
XX The invention relates to identifying compounds which are modulators
XX of syncytin expression. The syncytin modulators are useful in diagnosis
XX and treatment of preeclampsia and gestational trophoblast disorders (e.g.
XX choriocarcinoma, hydatiform mole, placental site tumour and missed/
XX incomplete abortion). Syncytin is a human gene derived from the
XX envelope gene of human endogenous defective retrovirus, HERV-W. The
XX present invention is based partly on the discovery that syncytin
XX expression is dramatically reduced in preeclampsia, and is also
XX mis-localised to the apical syncytiotrophoblast membrane. The present
XX sequence is human syncytin cDNA.
XX
XX Sequence 2930 BP; 842 A; 800 C; 571 G; 717 T; 0 other;
XX
XX Query Match 100.0%; Score 201; DB 24; Length 2930;
XX Best Local Similarity 100.0%; Pred. No. 6,1e-55;
XX Matches 201; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX 1 CCTACGACGAGTAGTCCTTCAAAATCGAAGAGCTTTAGACTTGTACCGCTGAAGAAGG 60
XX |
XX DB 2051 CCTACGACGAGTAGTCCTTCAAAATCGAAGAGCTTTAGACTTGTACCGCTGAAGAAGG 2110
XX |
XX QY 61 GGGAACTGTTATTTTAAAGGGAAGATGCTTATATATGTAATCAATCCGGAATGCT 120
XX |
XX DB 2111 GGGAACTGTTATTTTAAAGGGAAGATGCTTATATATGTAATCAATCCGGAATGCT 2170
XX |
XX QY 121 CACTGAGAAAGTTAAAGAAATTCGAGATCGAATACAGCTAGAGAGAGAGCTTCGAAA 180
XX |
XX DB 2171 CACTGAGAAAGTTAAAGAAATTCGAGATCGAATACAGCTAGAGAGAGAGCTTCGAAA 2230
XX |
XX QY 181 CACTGAGACCTGGGGCTCTCT 201
XX |
XX DB 2231 CACTGAGACCTGGGGCTCTCT 2251
XX
XX
XX RESULT 6
XX ID AAX77526 standard; cDNA; 2946 BP.
XX
XX AAX77526;
XX
XX 10-AUG-1999 (first entry)
XX

DE Human secreted protein AJ172_2 cDNA.
XX
XX Secreted protein; testes; brain; blood; placenta; human; murine; thymus;
XX bone marrow; treatment; prevention; nutrition; cytokine; immune; vaccine;
XX cell proliferation; cell differentiation; suppressor; tumour inhibition;
XX haematopoiesis regulator; activin; inhibin; chemotactic; chemokine;
XX haemostatic; thrombolytic; receptor; ligand; anti-inflammatory; tumour;
XX cadherin; tumour invasion suppressor; gene therapy; tissue growth; ss.
XX
XX Homo sapiens.
XX
XX WO9926972-A1.
XX
XX 03-JUN-1999.
XX
XX 17-NOV-1998; 98WO-US24614.
XX
XX 20-OCT-1998; 98US-0175928.
XX 21-NOV-1997; 97US-0976110.
XX 18-MAY-1998; 98US-0080478.
XX
XX (GENY) GENETICS INST INC.
XX
XX Collins-Racie LA, Evans C, Jacobs K, Lavallie ER;
XX McCoy JM, Merberg D, Treacy M;
XX WPI: 1999-357813/30.
XX P-PSDB; AAY08622.
XX
XX New polynucleotides encoding secreted proteins
XX
XX Claim 13a; Page 100-101; 142pp; English.
XX
XX This invention describes novel human secreted proteins encoded by
XX polynucleotides isolated from human adult testes, adult brain, adult
XX blood or adult placenta, or murine adult bone marrow or thymus cDNA
XX libraries. The products of the invention are predicted to have biological
XX activities which would make them suitable for treating, preventing or
XX ameliorating medical conditions in humans and animals, although no
XX supporting data is given. Suggested activities include nutritional
XX activity, cytokine and cell proliferation/differentiation activity,
XX immune stimulating (e.g. as vaccines) or suppressing activity,
XX haematopoiesis regulating activity, tissue growth activity,
XX activin/inhibin activity, chemotactic/chemokine activity, haemostatic
XX and thrombolytic activity, receptor/ligand activity, anti-inflammatory
XX activity, cadherin/tumour invasion suppressor activity, and tumour
XX inhibition activity. The polynucleotides are also stated to be useful
XX for gene therapy.
XX
XX Sequence 2946 BP; 858 A; 801 C; 570 G; 717 T; 0 other;
XX
XX Query Match 100.0%; Score 201; DB 20; Length 2946;
XX Best Local Similarity 100.0%; Pred. No. 6,1e-55;
XX Matches 201; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX 1 CCTACGACGAGTAGTCCTTCAAAATCGAAGAGCTTTAGACTTGTACCGCTGAAGAAGG 60
XX |
XX DB 2049 CCTACGACGAGTAGTCCTTCAAAATCGAAGAGCTTTAGACTTGTACCGCTGAAGAAGG 2108
XX |
XX QY 61 GGGAACTGTTATTTTAAAGGGAAGATGCTTATATATGTAATCAATCCGGAATGCT 120
XX |
XX DB 2109 GGGAACTGTTATTTTAAAGGGAAGATGCTTATATATGTAATCAATCCGGAATGCT 2168
XX |
XX QY 121 CACTGAGAAAGTTAAAGAAATTCGAGATCGAATACAGCTAGAGAGAGAGCTTCGAAA 180
XX |
XX DB 2169 CACTGAGAAAGTTAAAGAAATTCGAGATCGAATACAGCTAGAGAGAGAGCTTCGAAA 2228
XX |
XX QY 181 CACTGAGACCTGGGGCTCTCT 201
XX |
XX DB 2229 CACTGAGACCTGGGGCTCTCT 2249
XX
XX
XX RESULT 7

AA259468
ID AA259468 standard; cDNA; 2946 BP.
XX
AC AA259468;
XX
DT 11-APR-2000 (first entry)
XX
DE Human secreted protein AJ172_2 polynucleotide sequence.
XX
KW Human; secreted protein; disease diagnosis; pre-eclampsia; cancer;
KW placental pathology; metastasis inhibition; nutritional activity;
KW immune stimulator; haematopoiesis regulator; tissue growth;
KW tumour inhibitor; anti-inflammatory; clone AJ172_2; ATCC_98115;
KW gene therapy; ss.
XX
OS Homo sapiens.
XX
PN WO960020-A1.
XX
PD 25-NOV-1999.
XX
PF 17-MAY-1999; 99WO-US10915.
XX
PR 18-MAY-1998; 98US-0080478.
XX 20-OCT-1998; 98US-0175928.
XX
PA (GEMT) GENETICS INST INC.
XX
PI Jacobs K, McCoy JM, Lavallee ER, Collins-Racle LA, Evans C;
PI Merberg D, Ml S, Treacy M;
XX
DR WPI: 2000-116311/10.
DR P-PSDB: AA67313.
XX
PT New polynucleotides encoding secreted cDNA libraries, used to develop
PT products for the diagnosis and treatment of neoplastic disease
XX
PS Claim 14; Page 107-108; 149pp; English.
XX
CC This is the human secreted protein AJ172_2 nucleotide sequence, obtained
CC from a human adult testes cDNA library. The invention relates to secreted
CC human and murine proteins. The polynucleotides and proteins are predicted
CC to have biological activities which would make them suitable for
CC treating, preventing or ameliorating medical conditions in humans and
CC animals. Detection of the levels of the proteins can be used for the
CC diagnosis of e.g. pre-eclampsia, placental pathology or cancer. Agents
CC which modulate the expression or function of the proteins may be used for
CC treating a neoplastic disease and inhibiting metastasis. Other suggested
CC activities include nutritional activity (e.g. in feeds), cytokine and
CC cell proliferation/differentiation activity, immune stimulating
CC (e.g. as vaccines) or suppressing activity, haematopoiesis regulating
CC activity, tissue growth activity, activin/inhibin activity,
CC chemotactic/chemokinetic activity, haemostatic and thrombolytic activity,
CC receptor/ligand activity, anti-inflammatory activity, cadherin/tumour
CC invasion suppressor activity, and tumour inhibition activity. The
CC polynucleotide sequences are also stated to be useful for gene therapy.
XX
SQ Sequence 2946 BP; 858 A; 801 C; 570 G; 717 T; 0 other;
XX
Query Match 100.0%; Score 201; DB 21; Length 2946;
Best Local Similarity 100.0%; Pred. No. 6.1e-55;
Matches 201; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 2169 CACTGAGAAAGTTAAAGAAATTCGAGATCGAATACACGTAGACGAGAGCTTCGAAA 2228
QY 181 CACTGAGACCTGGGCGCTCCT 201
DB 2229 CACTGAGACCTGGGCGCTCCT 2249
XX
DE Human secreted protein AJ172_2 polynucleotide sequence.
XX
KW Human; secreted protein; disease diagnosis; pre-eclampsia; cancer;
KW placental pathology; metastasis inhibition; nutritional activity;
KW immune stimulator; haematopoiesis regulator; tissue growth;
KW tumour inhibitor; anti-inflammatory; clone AJ172_2; ATCC_98115;
KW gene therapy; ss.
XX
OS Homo sapiens.
XX
PN WO960020-A1.
XX
PD 25-NOV-1999.
XX
PF 17-MAY-1999; 99WO-US10915.
XX
PR 18-MAY-1998; 98US-0080478.
XX 20-OCT-1998; 98US-0175928.
XX
PA (INRM) INSERM INST NAT SANTE & RECH MEDICALE.
XX
PI Alliel PM, Perin J, Rieger F;
XX
DR WPI: 2000-160587/14.
XX
PT New nucleic acid sequences of human endogenous retroviruses, HERV-7q,
PT used for diagnosis, treatment and prevention of autoimmune and
PT neurological diseases
XX
PS Claim 3; Fig 1; 225pp; French.
XX
CC The present invention relates to new nucleic acid sequences of human
CC endogenous retroviruses, HERV-7q, which is located on chromosome 7q.
CC Regulatory elements associated with HERV-7q may alter expression of other
CC genes (even remote genes) on the same chromosome, inducing immunological
CC and/or neurological changes (which may be pathological or protective/
CC curative). HERV-7q peptides can be used to improve efficiency of the
CC immune response, e.g. in immunotherapy. HERV-7q peptides and their coding
CC sequences can be used in immunogenic or vaccinating compositions, for
CC protection against autoimmune diseases, particularly multiple sclerosis.
CC The peptides may also be used (by sequence comparison) to detect/identify
CC endogenous retroviruses that are abnormally expressed in cancer,
CC neuropathologies or other autoimmune diseases. The present sequence was
CC used to illustrate the invention.
XX
SQ Sequence 10499 BP; 3048 A; 2676 C; 2280 G; 2495 T; 0 other;
XX
Query Match 100.0%; Score 201; DB 21; Length 10499;
Best Local Similarity 100.0%; Pred. No. 9.7e-55;
Matches 201; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 181 CACTGACCCCTGGGCTCTCT 201
 |||||
 Db 9180 CACTGACCCCTGGGCTCTCT 9200

RESULT 9
 ABL61744
 ID ABL61744 standard; DNA; 56093 BP.
 XX

AC ABL61744;

DT 15-MAY-2002 (first entry)

DE Colon adenocarcinoma related gene sequence SEQ ID NO:81.

XX Human; cancer; colon; breast; ovary; oesophagus; kidney; thyroid;
 KW stomach; lung; prostate; pancreas; carcinoma; antitumour; cancerous;
 KW cytoskeletal; gene therapy; antineoplastic; Wilm's tumour; adenocarcinoma;
 KW gene; ds.
 XX

OS Homo sapiens.

XX WO200194629-A2.

XX 13-DEC-2001.

XX 30-MAY-2001; 2001MO-US10838.

XX 05-JUN-2000; 2000US-209473P.

XX 05-JUN-2000; 2000US-209531P.

XX 18-SEP-2000; 2000US-231333P.

XX 18-SEP-2000; 2000US-233617P.

XX 20-SEP-2000; 2000US-234099P.

XX 20-SEP-2000; 2000US-234034P.

XX 20-SEP-2000; 2000US-234052P.

XX 22-SEP-2000; 2000US-234509P.

XX 22-SEP-2000; 2000US-234567P.

XX 25-SEP-2000; 2000US-234923P.

XX 25-SEP-2000; 2000US-234924P.

XX 25-SEP-2000; 2000US-235077P.

XX 25-SEP-2000; 2000US-235082P.

XX 25-SEP-2000; 2000US-235134P.

XX 25-SEP-2000; 2000US-235280P.

XX 26-SEP-2000; 2000US-235637P.

XX 26-SEP-2000; 2000US-235638P.

XX 27-SEP-2000; 2000US-235711P.

XX 27-SEP-2000; 2000US-235720P.

XX 27-SEP-2000; 2000US-235840P.

XX 28-SEP-2000; 2000US-236028P.

XX 28-SEP-2000; 2000US-236032P.

XX 28-SEP-2000; 2000US-236033P.

XX 28-SEP-2000; 2000US-236034P.

XX 28-SEP-2000; 2000US-236109P.

XX 28-SEP-2000; 2000US-236111P.

XX 29-SEP-2000; 2000US-236842P.

XX 29-SEP-2000; 2000US-236891P.

XX 02-OCT-2000; 2000US-237172P.

XX 02-OCT-2000; 2000US-237173P.

XX 02-OCT-2000; 2000US-237278P.

XX 02-OCT-2000; 2000US-237294P.

XX 02-OCT-2000; 2000US-237295P.

XX 02-OCT-2000; 2000US-237316P.

XX 03-OCT-2000; 2000US-237425P.

XX 03-OCT-2000; 2000US-237596P.

XX 03-OCT-2000; 2000US-237604P.

XX 03-OCT-2000; 2000US-237606P.

XX 03-OCT-2000; 2000US-237608P.

XX 01-NOV-2000; 2000US-244867P.

XX 01-NOV-2000; 2000US-245084P.

XX (AVAL-) AVALON PHARM.

XX Young PE, Augustus M, Carter KC, Ebner R, Endress G, Horrigan S;
 PI Soppet DR, Weaver Z;
 XX WPI; 2002-168264/24.

XX Screening for anti-neoplastic agent involves exposing cells to a
 PT chemical agent to be tested for anti-neoplastic activity, and
 PT determining a change in expression of a gene of a signature gene set
 XX

PS Claim 1; SEQ ID 81; 44pp; English.

XX The present invention describes a method (M1) for screening for an
 CC anti-neoplastic agent. The method involves exposing cells to a chemical
 CC agent to be tested for anti-neoplastic activity, determining a change in
 CC expression of at least one gene (I) of a signature gene set, where (I)
 CC comprises a sequence (S) selected from 8447 sequences (given in ABL61664
 CC to ABL70110), or is at least 95% identical to (S), where a change in
 CC expression is indicative of anti-neoplastic activity. (I) has cytoskeletal
 CC activity and can be used in gene therapy. M1 can be used for screening
 CC an anti-neoplastic agent, and can be used for producing a product which
 CC is the data collected with respect to the anti-neoplastic agent as a
 CC result of M1, and the data is sufficient to convey the chemical
 CC structure and/or properties of the agent. M1 can be used in the
 CC treatment of cancer such as colon, breast, stomach, lung, thyroid,
 CC oesophageal, ovarian, kidney, prostate or pancreatic cancer.
 CC adenocarcinoma, carcinoma, clear cell cancer, infiltrating ductal cancer,
 CC infiltrating lobular cancer, squamous cell carcinoma, neuroendocrine
 CC carcinoma, papillary carcinoma and Wilm's tumour.

XX Sequence 56093 BP; 16164 A; 12346 C; 10702 G; 16881 T; 0 other;

XX Query Match 100.0%; Score 201; DB 24; Length 56093;

XX Best Local Similarity 100.0%; Pred. No. 1.8e-54; Matches 201; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCTAGCAGCAGTACTCTCTCAAAATCGAAGAGCTTTAGACTTGTCTAACCGCTGAAGAGG 60

Db 37000 CCTAGCAGCAGTACTCTCTCAAAATCGAAGAGCTTTAGACTTGTCTAACCGCTGAAGAGG 37059

QY 61 GGGAACTGTTATTTTAAAGGGAAGATGCTGTATATTTAATCAATCGGAATCGT 120

Db 37060 GGGAACTGTTATTTTAAAGGGAAGATGCTGTATATTTAATCAATCGGAATCGT 37119

QY 121 CACTGAGAACTTAAGAAATTCGATCGAATCAACGTAGACAGAGAGCTTCGAAA 180

Db 37120 CACTGAGAACTTAAGAAATTCGATCGAATCAACGTAGACAGAGAGCTTCGAAA 37179

QY 181 CACTGACCCCTGGGCTCTCT 201

Db 37180 CACTGACCCCTGGGCTCTCT 37200

RESULT 10

AAAX25665
 ID AAAX25665 standard; CDNA to mRNA; 7582 BP.

XX AAAX25665;

DT 21-MAY-1999 (first entry)

DE Complete human endogenous retrovirus W genome.

XX Clone; human endogenous retrovirus; genome; autoimmune disease;

KW multiple sclerosis; rheumatoid polyarthritis; insulin-dependent diabetes;

XX disseminated lupus erythematosus; pregnancy; chromosomal marker; ss.

OS Human endogenous retrovirus.

XX WO9902696-A1.

XX 21-JAN-1999.

```
PF 06-JUL-1998; 98WO-FR01442.
XX
PR 07-JUL-1997; 97FR-0008815.
XX
PA (INMR ) BIO MERIEUX.
XX
PI Beseme F, Blond JL, Boulton O, Mallet F, Mandrand B;
XX WPI; 1999-120897/10.
DR
XX New nucleic acid sequences from human endogenous retrovirus-W -
PT expressed exclusively in placenta and useful in diagnosis and
PT therapy of autoimmune disease, and abnormal or failed pregnancy
XX
PS Claim 1: Page 71-74; 106pp; French.
XX
XX This sequence represents the complete sequence of the human endogenous
CC retrovirus (HERV) W genome. The nucleic acids, their fragments or
CC peptides encoded by them are markers of autoimmune disease (e.g. multiple
CC sclerosis, rheumatoid polyarthritis, disseminated lupus erythematosus,
CC insulin- dependent diabetes and related pathologies) and of abnormal or
CC unsuccessful pregnancy and can be used as chromosomal markers for
CC susceptibility to these conditions, or proximately markers of genes
CC associated with this susceptibility.
XX
SQ Sequence 7582 BP; 2156 A; 1877 C; 1537 G; 1796 T; 2 U; 214 other;
Query Match 98.6%; Score 198.2; DB 20; Length 7582;
Best Local Similarity 96.5%; Pred. No. 6.9e-54;
Matches 194; Conservative 7; Mismatches 0; Indels 0; Gaps 0;
QY 1 CCTAGCAGCAGTAGTCTCTCAAAATCGAAGAGCTTTAGACTTGCTAACCGCTGAAGAGG 60
DB 6702 CCTAGCAGCAGTAGTCTCTCAAAATCGAAGAGCTTTAGACTTGCTAACCGCTGAAGAGG 6761
QY 61 GGGAACTGTTTATTTTAAAGGAGAGATGCTGTATTTATTCATATCCGAATCGT 120
DB 6762 GGGAACTGTTTATTTTAAAGGAGAGATGCTGTATTTATTCATATCCGAATCGT 6821
QY 121 CACTGAGAAAGTTAAAGAAATTCGAGATCGAATACAGCTAGAGAGAGCTTCGAAA 180
DB 6822 CACTGAGAAAGTTAAAGAAATTCGAGATCGAATACAGCTAGAGAGAGCTTCGAAA 6881
QY 181 CACTGAGACCTGGGCGCTCCT 201
DB 6882 CACTGAGACCTGGGCGCTCCT 6902
RESULT 11
AAA59215
ID AAA59215 standard; DNA; 7582 BP.
XX
AC AAA59215;
XX
XX 07-NOV-2000 (first entry)
XX Human endogenous retrovirus W (HERV-W) sequence.
XX
XX Autoimmune disease; retrovirus; human endogenous retrovirus W; HERV-W;
KW gag gene; pregnancy; multiple sclerosis; T cell proliferation; ss.
XX
XX Human endogenous retrovirus.
XX
XX Key Location/Qualifiers
FH 1..120
FT /*tag- a
FT /*note- "R of 5' LTR"
FT LTR 121..575
FT /*tag- b
FT /*note- "U5 of 5' LTR"
FT primer_bind 579..596
FT /*tag- c
FT CDS 5581..7164
```

```
FT /*tag- d
FT /*note- "ORF1 env538"
FT CDS 7039..7194
FT /*tag- e
FT /*note- "ORF2 52 AA"
FT CDS 7112..7255
FT /*tag- f
FT /*note- "ORF3 48 AA"
FT misc_feature 7244..7254
FT /*tag- g
FT /*note- "polypurine tract"
FT LTR 7256..7582
FT /*tag- h
FT /*note- "U3-R of 3' LTR"
FT polyA_signal 7563..7569
FT /*tag- i
XX
XX WO200043521-A2.
XX
XX 27-JUL-2000.
XX
XX 21-JAN-2000; 2000WO-FR00144.
XX
XX 21-JAN-1999; 99FR-0000888.
XX
XX (INMR ) BIO MERIEUX.
XX
XX Paranhos-Baccala G, Mallet F, Voisset C;
XX WPI; 2000-499229/44.
XX
XX New nucleic acid from human endogenous retrovirus, useful e.g. for
PT diagnosis of autoimmune disease and complications of pregnancy,
PT contains at least part of the gag gene
XX
XX Disclosure; Page 49-52; 53pp; French.
XX
XX The present sequence represents an endogenous retrovirus, which is
CC associated with an autoimmune disease, and is integrated into the human
CC genome. The retrovirus is human endogenous retrovirus W (HERV-W). The
CC HERV-W retrovirus is associated with autoimmune disease, failure of
CC pregnancy or disorders of pregnancy. HERV-W nucleic acid fragments, or
CC proteins derived from it, are useful for diagnosis of autoimmune
CC disease (specifically multiple sclerosis) and for monitoring pregnancy.
CC The nucleic acid fragments may also be used for in situ labelling of
CC isolated chromosomes, while the transcription product can be used to
CC study or monitor T cell proliferation in vitro.
XX
SQ Sequence 7582 BP; 2156 A; 1876 C; 1538 G; 1796 T; 216 other;
Query Match 98.6%; Score 198.2; DB 21; Length 7582;
Best Local Similarity 96.5%; Pred. No. 6.9e-54;
Matches 194; Conservative 7; Mismatches 0; Indels 0; Gaps 0;
QY 1 CCTAGCAGCAGTAGTCTCTCAAAATCGAAGAGCTTTAGACTTGCTAACCGCTGAAGAGG 60
DB 6702 CCTAGCAGCAGTAGTCTCTCAAAATCGAAGAGCTTTAGACTTGCTAACCGCTGAAGAGG 6761
QY 61 GGGAACTGTTTATTTTAAAGGAGAGATGCTGTATTTATTCATATCCGAATCGT 120
DB 6762 GGGAACTGTTTATTTTAAAGGAGAGATGCTGTATTTATTCATATCCGAATCGT 6821
QY 121 CACTGAGAAAGTTAAAGAAATTCGAGATCGAATACAGCTAGAGAGAGCTTCGAAA 180
DB 6822 CACTGAGAAAGTTAAAGAAATTCGAGATCGAATACAGCTAGAGAGAGCTTCGAAA 6881
QY 181 CACTGAGACCTGGGCGCTCCT 201
DB 6882 CACTGAGACCTGGGCGCTCCT 6902
RESULT 12
AAX25660
```

ID AAX25660 standard; cDNA to mRNA; 1136 BP.
XX AAX25660;
AC
XX 21-MAY-1999 (first entry)
DT
XX Human endogenous retrovirus W clone cl.C4C5.
DE
XX Clone: human endogenous retrovirus; genome; autoimmune disease;
KW multiple sclerosis; rheumatoid polyarthritis; insulin-dependent diabetes;
KM disseminated lupus erythematosus; pregnancy; chromosomal marker; ss.
XX
XX Human endogenous retrovirus.
OS
XX WO9902696-A1.
XX
XX 21-JAN-1999.
PD
XX 06-JUL-1998; 98WO-FR01442.
PF
XX 07-JUL-1997; 97FR-0008815.
PR
XX (INMR) BIO MERIEUX.
PA
XX Beseme F, Blond JL, Bouton O, Mallet F, Mandrand B;
PI
XX WPI; 1999-120897/10.
DR
XX
XX New nucleic acid sequences from human endogenous retrovirus-W -
PT expressed exclusively in placenta and useful in diagnosis and
PT therapy of autoimmune disease, and abnormal or failed pregnancy
XX
XX Claim 1; Page 59-60; 106pp; French.
PS
XX This sequence represents clone cl.C4C5 of the human endogenous retrovirus
CC (HERV) W genome. The nucleic acids, their fragments or peptides encoded
CC by them are markers of autoimmune disease (e.g. multiple sclerosis,
CC rheumatoid polyarthritis, disseminated lupus erythematosus, insulin-
CC dependent diabetes and related pathologies) and of abnormal or
CC unsuccessful pregnancy and can be used as chromosomal markers for
CC susceptibility to these conditions, or proximity markers of genes
CC associated with this susceptibility.
XX
XX Sequence 1136 BP; 336 A; 289 C; 241 G; 270 T; 0 other;
SQ
Query Match 98.4%; Score 197.8; DB 20; Length 1136;
Best Local Similarity 99.0%; Pred. No. 4.7e-54;
Matches 199; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
OY 1 CCTAGCAGCAGTAGTCCTTCAAAATCGAAGAGCTTTAGACTTGCTAACCGCTGAAGAAG 60
DB 236 CCTAGCAGCAGTAGTCCTTCAAAATCGAAGAGCTTTAGACTTGCTAACCGCTGAAGAAG 295
OY 61 GGGAACTGTTATTTTAAAGGGAAGATGCTGTTATTTATGTTATCAATCCGAAATGCT 120
DB 296 GGGAACTGTTATTTTAAAGGGAAGATGCTGTTATTTATGTTATCAATCCGAAATGCT 355
OY 121 CACTGAGAAAGTTAAAGAAATTCGAGATCGAATACAGCTGAGAGAGAGACTTCGAAA 180
DB 356 CACTGAGAAAGTTAAAGAAATTCGAGATCGAATACAGCTGAGAGAGAGACTTCGAAA 415
OY 181 CACTGAGACCTGGGGCTCTCT 201
DB 416 CACTGAGACCTGGGGCTCTCT 436
RESULT 13
AAS84210
ID AAS84210 standard; DNA; 1136 BP.
XX
XX AAS84210;
AC
XX 07-NOV-2000 (first entry)
DT

XX 3' pol gene and 3' non coding sequences of HERV-W from human genome.
DE
XX Autoimmune disease; retrovirus; human endogenous retrovirus W; HERV-W;
KW gag gene; pregnancy; multiple sclerosis; T cell proliferation; ss.
KM
XX Homo sapiens.
OS
XX WO200043521-A2.
XX
XX 27-JUL-2000.
PD
XX 21-JAN-2000; 2000WO-FR00144.
PF
XX 21-JAN-1999; 99FR-000888.
PR
XX (INMR) BIO MERIEUX.
PA
XX Paranhos-Baccala G, Mallet F, Volasset C;
PI
XX WPI; 2000-499229/44.
DR
XX
XX New nucleic acid from human endogenous retrovirus, useful e.g. for
PT diagnosis of autoimmune disease and complications of pregnancy,
PT contains at least part of the gag gene
XX
XX Disclosure; Page 46; 53pp; French.
PS
XX The present sequence represents an endogenous retroviral nucleic acid
CC fragment, which is associated with an autoimmune disease, and is
CC integrated into the human genome. The fragment is originally derived
CC from a novel retrovirus, human endogenous retrovirus W (HERV-W). The
CC HERV-W retrovirus is associated with autoimmune disease, failure of
CC pregnancy or disorders of pregnancy. The nucleic acid fragment, or
CC proteins derived from it, are useful for diagnosis of autoimmune
CC disease (specifically multiple sclerosis) and for monitoring pregnancy.
CC The nucleic acid fragments may also be used for in situ labelling of
CC isolated chromosomes, while the transcription product can be used to
CC study or monitor T cell proliferation in vitro.
XX
XX Sequence 1136 BP; 336 A; 289 C; 241 G; 270 T; 0 other;
SQ
Query Match 98.4%; Score 197.8; DB 21; Length 1136;
Best Local Similarity 99.0%; Pred. No. 4.7e-54;
Matches 199; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
OY 1 CCTAGCAGCAGTAGTCCTTCAAAATCGAAGAGCTTTAGACTTGCTAACCGCTGAAGAAG 60
DB 236 CCTAGCAGCAGTAGTCCTTCAAAATCGAAGAGCTTTAGACTTGCTAACCGCTGAAGAAG 295
OY 61 GGGAACTGTTATTTTAAAGGGAAGATGCTGTTATTTATGTTATCAATCCGAAATGCT 120
DB 296 GGGAACTGTTATTTTAAAGGGAAGATGCTGTTATTTATGTTATCAATCCGAAATGCT 355
OY 121 CACTGAGAAAGTTAAAGAAATTCGAGATCGAATACAGCTGAGAGAGAGACTTCGAAA 180
DB 356 CACTGAGAAAGTTAAAGAAATTCGAGATCGAATACAGCTGAGAGAGAGACTTCGAAA 415
OY 181 CACTGAGACCTGGGGCTCTCT 201
DB 416 CACTGAGACCTGGGGCTCTCT 436
RESULT 14
AAS84210
ID AAS84210 standard; cDNA; 6394 BP.
XX
XX AAS84210;
AC
XX 13-FEB-2002 (first entry)
DT
XX DNA encoding novel human diagnostic protein #20014.
DE
XX

KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
KM food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX
OS Homo sapiens.
XX
PN WO2001/5067-A2.
XX
PD 11-OCT-2001.
XX
PF 30-MAR-2001; 2001WO-US08631.
XX
PR 31-MAR-2000; 2000US-0540217.
PR 23-AUG-2000; 2000US-0649167.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Drmanac RT, Liu C, Tang YT;
XX
DR WPI; 2001-639362/73.
XX
P-PSDB; ABG20023.
XX
PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity
XX
PS Claim 1; SEQ ID No 20014; 103pp; English.
XX
XX The invention relates to isolated polynucleotide (I) and
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. AAs64197-AAS94564 represent novel human
CC diagnostic coding sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 6394 BP; 1840 A; 1597 C; 1384 G; 1571 T; 2 other;
Query Match 98.4%; Score 197.8; DB 23; Length 6394;
Best Local Similarity 99.0%; Pred. No. 8.8e-54;
Matches 199; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
OY 1 CCTAGCAGCAGTAGTCTCTTAATAATCGAAGAGCTTTAGACTGTACCCGCTGAAGAAG 60
DB 1157 CCTAGCAGCAGTAGTCTCTTAATAATCGAAGAGCTTTAGACTGTACCCGCTGAAGAAG 1216
OY 61 GGGAACTGTTTATTTTATAGGGAAGAATGCTGTATTATGTTAATCAATCCGGAATCGT 120
DB 1217 GGGAACTGTTTATTTTATAGGGAAGAATGCTGTATTATGTTAATCAATCCGGAATCGT 1276
OY 121 CACTGAGAAAGTTAAAGAAATTCGAGATCGAATCAACAGTAGAGAGAGAGCTTCGAAA 180
DB 1277 CACTGAGAAAGTTAAAGAAATTCGAGATCGAATCAACAGTAGAGAGAGAGCTTCGAAA 1336
OY 181 CACTGAGACCTGGGGCTCTCT 201
DB 1337 CACTGAGACCTGGGGCTCTCT 1357

RESULT 15
AAAX25661
ID AAAX25661 standard; cDNA to mRNA; 2782 BP.
XX
XX AAAX25661;
AC
XX
DT 21-MAY-1999 (first entry)
XX
DE Human endogenous retrovirus W clone cl.PH74.
XX
KW Clone; human endogenous retrovirus; genome; autoimmune disease;
KW multiple sclerosis; rheumatoid polyarthritis; insulin-dependent diabetes;
KW disseminated lupus erythematosus; pregnancy; chromosomal marker; ss.
XX
OS Human endogenous retrovirus.
XX
PN WO902696-A1.
XX
PD 21-JAN-1999.
XX
PF 06-JUL-1998; 98WO-FR01442.
XX
PR 07-JUL-1997; 97FR-0008815.
XX
PA (INMR) BIO MERIEUX.
XX
PI Beseme F, Blond JL, Bouton O, Mallet F, Mandrand B;
XX
DR WPI; 1999-120897/10.
XX
PT New nucleic acid sequences from human endogenous retrovirus-W -
PT expressed exclusively in placenta and useful in diagnosis and
PT therapy of autoimmune disease, and abnormal or failed pregnancy
XX
PS Claim 1; Page 60-63; 106pp; French.
XX
XX This sequence represents clone cl.PH74 of the human endogenous retrovirus
CC (HERV) W genome. The nucleic acids, their fragments or peptides encoded
CC by them are markers of autoimmune disease (e.g. multiple sclerosis,
CC rheumatoid polyarthritis, disseminated lupus erythematosus, insulin-
CC dependent diabetes and related pathologies) and of abnormal or
CC unsuccessful pregnancy and can be used as chromosomal markers for
CC susceptibility to these conditions, or proximity markers of genes
CC associated with this susceptibility.
XX
SQ Sequence 2782 BP; 741 A; 767 C; 565 G; 709 T; 0 other;
Query Match 96.0%; Score 193; DB 20; Length 2782;
Best Local Similarity 97.5%; Pred. No. 2.3e-52;
Matches 196; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
OY 1 CCTAGCAGCAGTAGTCTCTTAATAATCGAAGAGCTTTAGACTGTACCCGCTGAAGAAG 60
DB 1884 CCTAGCAGCAGTAGTCTCTTAATAATCGAAGAGCTTTAGACTGTACCCGCTGAAGAAG 1943
OY 61 GGGAACTGTTTATTTTATAGGGAAGAATGCTGTATTATGTTAATCAATCCGGAATCGT 120
DB 1944 GGGAACTGTTTATTTTATAGGGAAGAATGCTGTATTATGTTAATCAATCCGGAATCGT 2003
OY 121 CACTGAGAAAGTTAAAGAAATTCGAGATCGAATCAACAGTAGAGAGAGAGCTTCGAAA 180
DB 2004 CACTGAGAAAGTTAAAGAAATTCGAGATCGAATCAACAGTAGAGAGAGAGCTTCGAAA 2063
OY 181 CACTGAGACCTGGGGCTCTCT 201
DB 2064 CACTGAGACCTGGGGCTCTCT 2084

Search completed: May 2, 2003, 14:52:10
Job time : 185.333 secs

GenCore version 5.1.5
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 2, 2003, 14:19:30 ; Search time 1112 Seconds
(without alignments)
2927.422 Million cell updates/sec

Title: US-09-719-554-3_COPY_9000_9200

Perfect score: 201
Sequence: 1 ccctgacagcagtagctcctc.....actgacccctgggctcct 201

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 40 summaries

Database :

EST: *
1: em_estdb: *
2: em_estdb: *
3: em_estdb: *
4: em_estdb: *
5: em_estdb: *
6: em_estdb: *
7: em_estdb: *
8: em_estdb: *
9: em_estdb: *
10: em_estdb: *
11: em_estdb: *
12: em_estdb: *
13: em_estdb: *
14: em_estdb: *
15: em_estdb: *
16: em_estdb: *
17: em_estdb: *
18: em_estdb: *
19: em_estdb: *
20: em_estdb: *
21: em_estdb: *
22: em_estdb: *
23: em_estdb: *
24: em_estdb: *
25: em_estdb: *
26: em_estdb: *
27: em_estdb: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	201	100.0	322	12	BF990902 RC6-GN007
2	201	100.0	411	12	BG004254 RC6-GN007
3	201	100.0	494	14	BQ365207 RC6-GN007
4	201	100.0	714	12	BE734284 601565487
5	200	99.5	354	14	BQ365139 RC6-GN007
6	198.4	98.7	367	12	BF990707 RC6-GN007

c	7	197.8	98.4	346	14	BQ365143	BQ365143 RC6-GN007
c	8	187	93.0	702	9	AU138097	AU138097 AU138097
c	9	185	92.0	310	14	BQ365206	BQ365206 RC6-GN007
c	10	172	85.6	385	17	AQ062249	AQ062249 CTT-HSP-2
c	11	167	83.1	586	10	BE019603	BE019603 ba84f03.y
c	12	163.2	81.2	572	14	BQ365363	BQ365363 MR2-GN002
c	13	162.8	81.0	302	12	BG004247	BG004247 RC6-GN007
c	14	157.8	78.5	608	17	AQ072469	AQ072469 HS-5403.A
c	15	152.4	75.8	790	12	BG572445	BG572445 602593490
c	16	149.8	74.5	294	14	R01948	R01948 ye85b02.r1
c	17	143.8	71.5	723	9	AU138405	AU138405 AU138405
c	18	135.4	67.4	289	13	BI053219	BI053219 PMO-GN021
c	19	133.8	66.6	316	13	BI053220	BI053220 PMO-GN021
c	20	129.8	64.6	305	12	BF990710	BF990710 RC6-GN007
c	21	127	63.2	427	9	AA776439	AA776439 zj30h06.s
c	22	125	62.2	433	14	N77302	N77302 yv43e03.r1
c	23	122.4	60.9	449	14	N58614	N58614 yv56h11.r1
c	24	106.6	53.0	385	14	T93615	T93615 ye05g09.s1
c	25	102	50.7	666	13	BG911940	BG911940 602809647
c	26	99.4	49.5	665	17	AG066067	AG066067 Pan trogl
c	27	98.8	49.2	323	14	R00744	R00744 ye74h09.s1
c	28	93	46.3	422	17	AQ150723	AQ150723 HS-3203.A
c	29	82	40.8	264	17	AQ151165	AQ151165 HS-2186.A
c	30	79.8	39.7	314	9	AA960921	AA960921 oos1a07.s
c	31	77	38.3	771	13	BI087886	BI087886 602852690
c	32	76.8	38.2	314	17	AQ059232	AQ059232 RPT-11-3
c	33	74.4	37.0	179	14	T93945	T93945 ye42e02.r1
c	34	73.4	36.5	664	17	AG090771	AG090771 Pan trogl
c	35	71.8	35.7	471	10	AM847812	AM847812 IL3-CT021
c	36	71.4	35.7	471	10	AM851794	AM851794 OVO-CT022
c	37	70	34.8	634	10	BB635247	BB635247 B635247
c	38	70	34.8	710	17	B69218	B69218 CTT-HSP-205
c	39	68	33.8	636	17	AG116240	AG116240 Pan trogl
c	40	67.2	33.4	754	17	AG031563	AG031563 Pan trogl
c	41	67	33.3	374	14	BQ346732	BQ346732 IL5-NF007
c	42	66.2	32.9	370	12	BG011064	BG011064 OVI-GN031
c	43	65.8	32.7	414	9	AT508499	AT508499 vb34d08.y
c	44	65.8	32.7	425	17	AQ058380	AQ058380 HS-2068.B
c	45	65.4	32.5	361	13	BI977020	BI977020 486270 MA

ALIGNMENTS

RESULT 1
LOCUS BF990902 322 bp mRNA linear EST 23-JAN-2001
DEFINITION RC6-GN0070-271000-023-H12 GN0070 Homo sapiens CDNA, mRNA sequence.
ACCESSION BF990902
VERSION BF990902.1 GI:12397227
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
AUTHORS
1 (bases 1 to 322)
Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Brites, M.R.,
Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F.,
Goldman, G.H., Carvalho, A.F., Matsukuma, A., Bata, G.S., Simpson, D.H.,
Brunstein, A., de Oliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare,
M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and
Simpson, A.J.

TITLE
Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
JOURNAL
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
MEDLINE
20202663
COMMENT
Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001

Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=RC6&t2=RC6-GN0070-271000-023-H12&t3=2000-10-27&t4=1)
Seq primer: puc 18 forward
High quality sequence start: 22
High quality sequence stop: 275.

FEATURES

Source

Location/Qualifiers

1..322

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone_id="GN0070"

/dev_stage="Adult"

/note="Organ: placenta.normal; Vector: puc18; Site_1: SmaI; Site_2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent Application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the puc 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

BASE COUNT

95 a

77 c

72 g

78 t

Query Match

Best Local Similarity 100.0%; Score 201; DB 12; Length 322;

Matches 201; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

ORIGIN

QY 1 CCTAGCAGCAGTACTCCTTCAAAATCGAAGAGCTTGTAGACTGTGTAACCGCTGAAGAAG 60
|||||
DB 88 CCTAGCAGCAGTACTCCTTCAAAATCGAAGAGCTTGTAGACTGTGTAACCGCTGAAGAAG 147
|||||
QY 61 GGGAACTGTTATTTTATAGGGAAGAATGCTGTATATATCTTAATCAATCCGAATCGT 120
|||||
DB 148 GGGAACTGTTATTTTATAGGGAAGAATGCTGTATATATCTTAATCAATCCGAATCGT 207
|||||
QY 121 CACTGAGAAAGTTAAAGAAATTCGAGATCGAATCAACGATGAGACGAGGAGCTTCGAAA 180
|||||
DB 208 CACTGAGAAAGTTAAAGAAATTCGAGATCGAATCAACGATGAGACGAGGAGCTTCGAAA 267
|||||
QY 181 CACTGAGACCTGGGGCCTCCT 201
|||||
DB 268 CACTGAGACCTGGGGCCTCCT 288
|||||

RESULT 2

LOCUS

BG004254/c 411 bp mRNA linear EST 24-JAN-2001

DEFINITION

RC6-GN0070-301100-014-C06 GN0070 Homo sapiens cDNA, mRNA sequence.

ACCESSION

BG004254

VERSION

BG004254.1 GI:12445242

KEYWORDS

EST.

SOURCE

human.

ORGANISM

Homo sapiens

REFERENCE

1 (bases 1 to 411)

AUTHORS

Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R., Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F., Goldman,G.H., Carvalho,A.F., Matsukuma,A., Bala,G.S., Simpson,D.H., Brunstein,A., deoliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare

TITLE

Shotgun sequencing of the human transcriptome with ORF expressed

JOURNAL

Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)

MEDLINE

20202663

COMMENT

Contact: Simpson A.J.G. Laboratory of Cancer Genetics Ludwig Institute for Cancer Research Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil

Tel: +55-11-2704922.

Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=RC6&t2=RC6-GN0070-301100-014-C06&t3=2000-11-30&t4=1)
Seq primer: puc 18 forward
High quality sequence start: 21
High quality sequence stop: 411.

FEATURES

Source

Location/Qualifiers

1..411

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone_id="GN0070"

/dev_stage="Adult"

/note="Organ: placenta.normal; Vector: puc18; Site_1: SmaI; Site_2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent Application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the puc 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

BASE COUNT

103 a

94 c

96 g

118 t

Query Match

Best Local Similarity 100.0%; Score 201; DB 12; Length 411;

Matches 201; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

ORIGIN

QY 1 CCTAGCAGCAGTACTCCTTCAAAATCGAAGAGCTTGTAGACTGTGTAACCGCTGAAGAAG 60
|||||
DB 294 CCTAGCAGCAGTACTCCTTCAAAATCGAAGAGCTTGTAGACTGTGTAACCGCTGAAGAAG 235
|||||
QY 61 GGGAACTGTTATTTTATAGGGAAGAATGCTGTATATATCTTAATCAATCCGAATCGT 120
|||||
DB 234 GGGAACTGTTATTTTATAGGGAAGAATGCTGTATATATCTTAATCAATCCGAATCGT 175
|||||
QY 121 CACTGAGAAAGTTAAAGAAATTCGAGATCGAATCAACGATGAGACGAGGAGCTTCGAAA 180
|||||
DB 174 CACTGAGAAAGTTAAAGAAATTCGAGATCGAATCAACGATGAGACGAGGAGCTTCGAAA 115
|||||
QY 181 CACTGAGACCTGGGGCCTCCT 201
|||||
DB 114 CACTGAGACCTGGGGCCTCCT 94
|||||

RESULT 3

LOCUS

B0365207/c 494 bp mRNA linear EST 21-MAY-2002

DEFINITION

RC6-GN0070-220800-021-D12 GN0070 Homo sapiens cDNA, mRNA sequence.

ACCESSION

B0365207

VERSION

B0365207.1 GI:21040719

KEYWORDS

EST.

SOURCE

human.

ORGANISM

Homo sapiens

REFERENCE

1 (bases 1 to 494)

AUTHORS

Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R., Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F., Goldman,G.H., Carvalho,A.F., Matsukuma,A., Bala,G.S., Simpson,D.H., Brunstein,A., deoliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare

TITLE

Shotgun sequencing of the human transcriptome with ORF expressed

JOURNAL

Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)

MEDLINE

20202663

COMMENT

Contact: Simpson A.J.G. Laboratory of Cancer Genetics Ludwig Institute for Cancer Research Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil

Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL
(<http://www.ludwig.org.br/scripts/gethtml2.pl?cl=RC6&cl2=RC6-GN0070-220800-021-D12&cl3=2000-08-22&cl4=1>)
Seq primer: puc 18 forward
High quality sequence start: 40
High quality sequence stop: 432.
Location/Qualifiers

FEATURES

source

1. .494
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="GN0070"
/dev_stage="Adult"
/note="Organ: placenta_normal; Vector: puc18; Site_1: Sma1
; Site_2: Sma1; A mini-library was made by cloning
products derived from ORESTES PCR (U.S. Letters Patent
Application No. 196,716 - Ludwig Institute for Cancer
Research) profiles into the pUC 18 vector. Reverse
transcription of tissue mRNA and cDNA amplification were
performed under low stringency conditions."

BASE COUNT

125 a 105 c 121 g 143 t

ORIGIN

Query Match

Best Local Similarity 100.0%; Score 201; DB 14; Length 494;
Matches 201; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCTAGCAGCAGTAGTCCTTCAAAATCGAAGAGCTTACCTGCTACCGCTGAAGAGG 60
DB 400 CCTAGCAGCAGTAGTCCTTCAAAATCGAAGAGCTTACCTGCTACCGCTGAAGAGG 341
QY 61 GGGAACTGTTATTTTATTTAGGGGAGAAATGCTGTTATTTATTCATCCGGAATCGT 120
DB 340 GGGAACTGTTATTTTATTTAGGGGAGAAATGCTGTTATTTATTCATCCGGAATCGT 281
QY 121 CACTGAGAAAGTTAAAGAAATTCGAGATCGAATACAGCTAGAGAGAGACTTCGAAA 180
DB 280 CACTGAGAAAGTTAAAGAAATTCGAGATCGAATACAGCTAGAGAGAGACTTCGAAA 221
QY 181 CACTGGACCTGGGGCCTCCT 201
DB 220 CACTGGACCTGGGGCCTCCT 200

RESULT 4

BE734284

LOCUS 601565487F1 NIH_MGC_21 Homo sapiens cDNA clone IMAGE:3840572 5',
DEFINITION mRNA sequence.

ACCESSION BE734284

VERSION BE734284.1 GI:10148276

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 714)
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.

AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.

TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL Unpublished (1999)

COMMENT Contact: Robert Strausberg, Ph.D.

Email: cgabs-remail.nih.gov

Tissue Procurement: ATCC

cDNA Library Preparation: Ling Hong/Rubin Laboratory

DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov

Plate: LICM530 row: m column: 21

High quality sequence stop: 712.

FEATURES

source

Location/Qualifiers

1. .714
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="IMAGE:3840572"
/clone_lib="NIH MGC 21"
/tissue_type="choriocarcinoma"
/lab_host="DH10B (phage-resistant)"
/note="Organ: placenta; Vector: pOTB7; Site_1: XhoI;
Site_2: EcoRI; cDNA made by oligo-dT priming.
Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGCACGAG(G). Size-selected >500bp
for average insert size 1.8kb. Library constructed by
Ling Hong in the laboratory of Gerald M. Rubin (University
of California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies)."

BASE COUNT

194 a 183 c 137 g 200 t

ORIGIN

Query Match 100.0%; Score 201; DB 12; Length 714;
Best Local Similarity 100.0%; Pred. No. 2.6e-51;
Matches 201; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCTAGCAGCAGTAGTCCTTCAAAATCGAAGAGCTTACCTGCTACCGCTGAAGAGG 60
DB 397 CCTAGCAGCAGTAGTCCTTCAAAATCGAAGAGCTTACCTGCTACCGCTGAAGAGG 456
QY 61 GGGAACTGTTATTTTATTTAGGGGAGAAATGCTGTTATTTATTCATCCGGAATCGT 120
DB 457 GGGAACTGTTATTTTATTTAGGGGAGAAATGCTGTTATTTATTCATCCGGAATCGT 516
QY 121 CACTGAGAAAGTTAAAGAAATTCGAGATCGAATACAGCTAGAGAGAGACTTCGAAA 180
DB 517 CACTGAGAAAGTTAAAGAAATTCGAGATCGAATACAGCTAGAGAGAGACTTCGAAA 576
QY 181 CACTGGACCTGGGGCCTCCT 201
DB 577 CACTGGACCTGGGGCCTCCT 597

RESULT 5

BQ365139

LOCUS RC6-GN0070-170800-011-D03 GN0070 Homo sapiens cDNA, mRNA sequence.
DEFINITION BQ365139

ACCESSION BQ365139.1 GI:21040651

VERSION BQ365139.1

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 354)

Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.

AUTHORS Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R.,
Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F.,
Goldman, G.H., Carvalho, A.F., Matsukuma, A., Bais, G.S., Simpson, D.H.,
Brunstein, A., deoliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare,
M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and
Simpson, A.J.

TITLE Shotgun sequencing of the human transcriptome with ORF expressed

JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)

MEDLINE 20020653

COMMENT Contact: Simpson A.J.G.

Laboratory of Cancer Genetics

Ludwig Institute for Cancer Research

Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil

Tel: +55-11-2704922

Fax: +55-11-2707001

Email: asimpson@ludwig.org.br

This sequence was derived from the FAPESP/LICR Human Cancer Genome

Project. This entry can be seen in the following URL

(<http://www.ludwig.org.br/scripts/gethtml2.pl?cl=RC6&cl2=RC6-GN0070->

170800-011-D03&f3=2000-08-17&f4=1)

Seq primer: puc 18 forward
High quality sequence start: 16
High quality sequence stop: 354.
Location/Qualifiers

FEATURES

SOURCE

1. 354

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone_id="GN0070"

/dev_stage="Adult"

/note="Organ: placenta, normal; Vector: puc18; Site_1: SmaI
; Site_2: SmaI; A mini-library was made by cloning
products derived from ORESTES PCR (U.S. Letters Patent
Application No. 196,716 - Ludwig Institute for Cancer
Research) profiles into the puc 18 vector. Reverse
transcription of tissue mRNA and cDNA amplification were
performed under low stringency conditions."

BASE COUNT 102 a 86 c 79 g 86 t 1 others
ORIGIN

Query Match 99.5%; Score 200; DB 14; Length 354;
Best Local Similarity 99.5%; Pred. No. 4.6e-51;
Matches 200; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 CCTAGCAGCAGTACTCTTCAAAATCGAAGAGCTTTAGACTTGTACCGCTGAAGAAG 60
|||
Db 120 CCTAGCAGCAGTACTCTTCAAAATCGAAGAGCTTTAGACTTGTACCGCTGAAGAAG 179
|||
OY 61 GGAACCTGTTATTTTATAGGGAAGAATGCTGTTATTTATTTATCAATCCGAATCGT 120
|||
Db 180 GGAACCTGTTATTTTATAGGGAAGAATGCTGTTATTTATTTATCAATCCGAATCGT 239
|||
OY 121 CACTGAGAAAGTTAAAGAAATTCGAGATCGAATCAACAGTAGAGAGAGAGCTTCGAAA 180
|||
Db 240 CACTGAGAAAGTTAAAGAAATTCGAGATCGAATCAACAGTAGAGAGAGAGCTTCGAAA 299
|||
OY 181 CACTGAGACCTTGGGGCCTCCT 201
|||
Db 300 CACTGAGACCTTGGGGCCTCCT 320
|||

RESULT 6
BF990707 367 bp mRNA linear EST 23-JAN-2001
LOCUS
DEFINITION RC6-GN0070-271000-013-D01 GN0070 Homo sapiens cDNA, mRNA sequence.
ACCESSION
VERSION BF990707.1 GI:12397032
KEYWORDS
SOURCE EST.
ORGANISM human.
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 367)

REFERENCE
AUTHORS
Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R.,
Nagal, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F.,
Goldman, G.H., Carvalho, A.F., Matsukuma, A., Bala, G.S., Simpson, D.H.,
Brunstein, A., deoliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare
M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and
Simpson, A.J.

TITLE
Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
JOURNAL
MEDLINE
COMMENT
20202663
Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL

(http://www.ludwig.org.br/scripts/gethtml2.pl?l1=RC6&f2=RC6-GN0070-
271000-013-D01&f3=2000-10-27&f4=1)

Seq primer: puc 18 forward
High quality sequence start: 17
High quality sequence stop: 359.
Location/Qualifiers

FEATURES

SOURCE

1. 367

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone_id="GN0070"

/dev_stage="Adult"

/note="Organ: placenta, normal; Vector: puc18; Site_1: SmaI
; Site_2: SmaI; A mini-library was made by cloning
products derived from ORESTES PCR (U.S. Letters Patent
Application No. 196,716 - Ludwig Institute for Cancer
Research) profiles into the puc 18 vector. Reverse
transcription of tissue mRNA and cDNA amplification were
performed under low stringency conditions."

BASE COUNT 104 a 88 c 85 g 89 t 1 others
ORIGIN

Query Match 98.7%; Score 198.4; DB 12; Length 367;
Best Local Similarity 99.0%; Pred. No. 1.4e-50;
Matches 199; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 CCTAGCAGCAGTACTCTTCAAAATCGAAGAGCTTTAGACTTGTACCGCTGAAGAAG 60
|||
Db 133 CCTAGCAGCAGCAGTACTCTTCAAAATCGAAGAGCTTTAGACTTGTACCGCTGAAGAAG 192
|||
OY 61 GGAACCTGTTATTTTATAGGGAAGAATGCTGTTATTTATTTATCAATCCGAATCGT 120
|||
Db 193 GGAACCTGTTATTTTATAGGGAAGAATGCTGTTATTTATTTATCAATCCGAATCGT 252
|||
OY 121 CACTGAGAAAGTTAAAGAAATTCGAGATCGAATCAACAGTAGAGAGAGAGCTTCGAAA 180
|||
Db 253 CACTGAGAAAGTTAAAGAAATTCGAGATCGAATCAACAGTAGAGAGAGAGCTTCGAAA 312
|||
OY 181 CACTGAGACCTTGGGGCCTCCT 201
|||
Db 313 CACTGAGACCTTGGGGCCTCCT 333
|||

RESULT 7
BQ365143/c 346 bp mRNA linear EST 21-MAY-2002
LOCUS
DEFINITION RC6-GN0070-170800-011-F12 GN0070 Homo sapiens cDNA, mRNA sequence.
ACCESSION
VERSION BQ365143
KEYWORDS
SOURCE EST.
ORGANISM human.
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 346)

REFERENCE
AUTHORS
Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R.,
Nagal, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F.,
Goldman, G.H., Carvalho, A.F., Matsukuma, A., Bala, G.S., Simpson, D.H.,
Brunstein, A., deoliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare
M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and
Simpson, A.J.

TITLE
Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
JOURNAL
MEDLINE
COMMENT
20202663
Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome

Project. This entry can be seen in the following URL:
(http://www.ludwig.org.br/scripts/gethtml2.pl?l=RC6&t2=RC6-GN0070-170800-011-F1&t3=2000-08-17&t4=1)

Seq primer: puc 18 forward
High quality sequence start: 18
High quality sequence stop: 346.
Location/Qualifiers

FEATURES

source 1..346

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="GN0070"
/dev_stage="Adult"
/note="Organ: placenta.normal; Vector: puc18; Site_1: SmaI
; Site_2: SmaI; A mini-library was made by cloning
products derived from OREBES PCR (U.S. Letters Patent
Application No. 196,716 - Ludwig Institute for Cancer
Research) profiles into the pUC 18 vector. Reverse
transcription of tissue mRNA and cDNA amplification were
performed under low stringency conditions."

BASE COUNT

84 a 84 c 79 g 99 t

ORIGIN

Query Match

Best Local Similarity 98.4%; Score 197.8; DB 14; Length 346;
Matches 199; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY

1 CCTAGACGACGATGCTCTTCAAAATCGAGAGCTTTAGACTTCTAACCCCTAAAGAGG 60

229 CCTAGACGACGATGCTCTTCAAAATCGAGAGCTTTAGACTTCTAACCCCTAAAGAGG 170

QY

61 GGGACCTGTTTATTTTGGGAGAGAGATGCTTTATTTATTAATCAATCCGGAATCGT 120

169 GGGAGCTGTTTATTTTGGGAGAGATGCTTTATTTATTAATCAATCCGGAATCGT 110

QY

121 CACTGAGAAAGTTAAAGAAATTCGAGATCGAATATCAACGTCAGAGAGAGCTTCGAAA 180

109 CACTGAGAAAGTTAAAGAAATTCGAGATCGAATATCAACGTCAGAGAGAGCTTCGAAA 50

QY

181 CACTGACCTGGGGCTCTCT 201

Db

49 CACTGACCTGGGGCTCTCT 29

RESULT 8

LOCUS AUI38097 702 bp mRNA linear EST 02-AUG-2002

DEFINITION

AUI38097 PLACEL Homo sapiens cDNA clone PLACE1007839 5', mRNA

ACCESSION

AUI38097

VERSION

AUI38097

KEYWORDS

EST

SOURCE

human.

ORGANISM

Homo sapiens

REFERENCE

1 (bases 1 to 702)

AUTHORS

Ota, T., Nishikawa, T., Suzuki, Y., Ishii, S., Saito, K., Kawai, Y., Yamamoto, J., Wakamatsu, A., Nakamura, Y., Nagai, T., Sugano, S. and Isogai, T.

TITLE

HRI human cDNA project

JOURNAL

Unpublished (2000)

COMMENT

Contact: Takao Isogai

FEATURES

Genomics Laboratory
Helix Research Institute
1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
Tel: 81-438-52-3975
Fax: 81-438-52-3986
Email: genomics@hri.co.jp
HRI human cDNA project; 5'- & 3'-end one pass sequencing; Helix
Research Institute; cDNA library construction; Department of
Virology, Institute of Medical Science, University of Tokyo, and
Helix Research Institute.
Location/Qualifiers

source

1..702

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="PLACE1007839"
/clone_lib="PLACE1"
/tissue_type="placenta"
/note="Vector: pME18SFL3"

BASE COUNT

206 a 173 c 131 g 189 t 3 others

ORIGIN

Query Match

Best Local Similarity 98.0%; Score 187; DB 9; Length 702;
Matches 188; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

QY

1 CCTAGACGACGATGCTCTTCAAAATCGAGAGCTTTAGACTTCTAACCCCTAAAGAGG 60

490 CCTAGACGACGATGCTCTTCAAAATCGAGAGCTTTAGACTTCTAACCCCTAAAGAGG 549

QY

61 GGGACCTGTTTATTTTGGGAGAGAGATGCTTTATTTATTAATCAATCCGGAATCGT 120

550 GGGAGCTGTTTATTTTGGGAGAGAGATGCTTTATTTATTAATCAATCCGGAATCGT 609

QY

121 CACTGAGAAAGTTAAAGAAATTCGAGATCGAATATCAACGTCAGAGAGAGCTTCGAAA 179

610 CACTGAGAAAGTTAAAGAAATTCGAGATCGAATATCAACGTCAGAGAGAGCTTCGAAA 669

QY

180 ACACGACCTGGGGCTCTCT 201

Db

670 ACACGACCTGGGGCTCTCT 691

RESULT 9

LOCUS BQ365206/c 310 bp mRNA linear EST 21-MAY-2002

DEFINITION BQ365206 BQ365206 Homo sapiens cDNA, mRNA sequence.

ACCESSION BQ365206.1 GI:21040718

VERSION BQ365206.1 GI:21040718

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.

1 (bases 1 to 310)

Dias Neto, E., Garcia Correa, R., Verjovsky-Almeida, S., Briones, M.R.,

Nagai, M.A., da Silva, M. Jr., Zago, M.A., Bordin, S., Costa, F.F.,

Goldman, G.H., Carvalho, A.F., Matsukuma, A., Baia, G.S., Simpson, D.H.,

Brunstein, A., de Oliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare

, M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and

Simpson, A.J.

Shotgun sequencing of the human transcriptome with ORF expressed

sequence tags

Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)

20202663

Contact: Simpson A.J.G.

Laboratory of Cancer Genetics

Ludwig Institute for Cancer Research

Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,

Brazil

Tel: +55-11-2704922

Fax: +55-11-2707001

Email: asimpson@ludwig.org.br

This sequence was derived from the FAPESP/LICR Human Cancer Genome

Project. This entry can be seen in the following URL

(http://www.ludwig.org.br/scripts/gethtml2.pl?l=RC6&t2=RC6-GN0070-220800-021-D1&t3=2000-08-22&t4=1)

Seq primer: puc 18 forward
High quality sequence start: 19
High quality sequence stop: 310.
Location/Qualifiers

FEATURES

source 1..310

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="GN0070"

```
/dev_stage="Adult"
/Note="Organ: Placenta;normal; Vector: puc18; Site_1: SmaI
; Site_2: SmaI; A mini-library was made by cloning
products derived from ORESSES PCR (U.S. Letters Patent
Application No. 196,716 - Ludwig Institute for Cancer
Research) profiles into the pUC 18 vector. Reverse
transcription of tissue mRNA and cDNA amplification were
performed under low stringency conditions."
```

BASE COUNT 74 a 73 c 68 g 95 t

ORIGIN

Query Match 92.0%; Score 185; DB 14; Length 310;
Best Local Similarity 97.4%; Pred. No. 1.9e-46;
Matches 188; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 CCTACAGACGATGCTTCCTCAAAATCGAAGACCTTAGACTGCTTACCGCTGGAAGAG 60
|||||
DB 193 CCTACAGACGATGCTTCCTCAAAATCGAAGACCTTAGACTGCTTACCGCTGGAAGAG 134
|||||
QY 61 GGGAACTGTATTATTTTGGGGAAGATGCTTATTATTTATTCATCCGGAATCGT 120
|||||
DB 133 GGGAACTGTATTATTTTGGGGAAGATGCTTATTATTTATTCATCCGGAATCGT 74
|||||
QY 121 CACTGAGAAAGTTAAAGAAATTCGAGATCGAATACACGTAGACGAGAGACCTTCGAA 180
|||||
DB 73 CACTGAGAAAGTTAAAGAAATTCGAGATCGAATACACGTAGACGAGAGACCTTCGAA 14
|||||
QY 181 CACTGAGACCTCGG 193
|||||
DB 13 CACTGAGACCTAGG 1

RESULT 10
A0062249/c 385 bp DNA linear GSS 31-JUL-1998
LOCUS CIT-HSP-2347N16.TR CIT-HSP Homo sapiens genomic clone 2347N16, DNA
DEFINITION sequence.
ACCESSION A0062249
VERSION A0062249.1 GI:3364161
KEYWORDS GSS.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 385)
Adams M.D., Rounsley S.D., Zhao S., Field C.E., Bass S., Linher K.,
Golden K., Berry K., Granger D., Suh E., Wible C., Shizuya H.,
Simon M., and Venter J.C. Use of a random BAC End Sequence Database for Sequence-Ready Map
Building (1998)
Unpublished (1998)
JOURNAL Other GSS: CIT-HSP-2347N16.TF
COMMENT Contact: Mark Adams
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: mdamas@tigr.org
Clones are available from Research Genetics (Info@resgen.com). BAC
end search page:
http://www.tigr.org/tdb/hungen/bac_end_search/bac_end_search.html.
Seq primer: M13 Reverse
Class: BAC ends.

FEATURES
source location/Qualifiers
1..385
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="2347N16"
/clone_lib="CIT-HSP"
/sex="Male"
/cell_type="Sperm"
/note="Vector: pBelOBAC11; Site_1: HindIII; Site_2:

```
BASE COUNT 109 a 80 c 91 g 105 t  
ORIGIN HindIII

Query Match 85.6%; Score 172; DB 17; Length 385;  
Best Local Similarity 93.6%; Pred. No. 2e-42;  
Matches 191; Conservative 0; Mismatches 10; Indels 3; Gaps 1;



QY 1 CCTACAGACGATGCTTCCTCAAAATCGAAGACCTTAGACTGCTTACCGCTGGAAG 57  
|||||  
DB 334 CCTACAGACGATGCTTCCTCAAAATCGAAGACCTTAGACTGCTTACCGCTGGAAG 275  
|||||  
QY 58 AGGGGAACCTGTATTATTTTGGGGAAGATGCTTATTATTTATTCATCCGGAAT 117  
|||||  
DB 274 AGGGGAACCTGTATTATTTTGGGGAAGATGCTTATTATTTATTCATCCGGAAT 215  
|||||  
QY 118 CACTGAGAAAGTTAAAGAAATTCGAGATCGAATACACGTAGACGAGAGACCTTCG 177  
|||||  
DB 214 CACTGAGAAAGTTAAAGAAATTCGAGATCGAATACACGTAGACGAGAGACCTTCG 155  
|||||  
QY 178 AACACTGAGACCTGGGCGCTCT 201  
|||||  
DB 154 AACACTGAGACCTGGGCGCTCT 131  
|||||


```

RESULT 11
BE019603 586 bp mRNA linear EST 06-JUN-2000
LOCUS ba04103.y1 NIH_MGC_21 Homo sapiens cDNA clone IMAGE:2907101 5'
DEFINITION similar to TR:095244 095244 ENVELOPE PROTEIN; contains Alu
repetitive element; mRNA sequence.
ACCESSION BE019603
VERSION BE019603.1 GI:8279682
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 586)
NIH-MGC http://mgi.nci.nih.gov/
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cga@bbs.femail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Ling Hong/Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNLN at:
image.llnl.gov/image/html/resources.shtml
Seq primer: -40RP from Gibco
High quality sequence stop: 519.

FEATURES
source location/Qualifiers
1..586
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:2907101"
/clone_lib="NIH_MGC_21"
/tissue_type="choriocarcinoma"
/lab_host="DH10B (phage-resistant)"
/note="Organ: Placenta; Vector: pOTB7; Site_1: XhoI;
Site_2: EcoRI; cDNA made by oligo-dT priming.
directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGCAGAG(C) Size-selected >500bp
for average insert size 1.8kb. Library constructed by
Ling Hong in the laboratory of Gerald M. Rubin (University
of California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies)."

BASE COUNT 168 a 141 c 113 g 164 t

ORIGIN

Query Match 83.1%; Score 167; DB 10; Length 586;

Best Local Similarity 100.0%; Pred. No. 7.5e-41;
Matches 167; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTTAGCAGCAGTACTCTTCAAAATCGAAGAGCTTTAG-ACCTGCTAACCGCTGAAGAG 60
|||||
Db 420 CTTAGCAGCAGTACTCTTCAAAATCGAAGAGCTTTAG-ACCTGCTAACCGCTGAAGAG 479
GGGAACTGTTTATTTTGGGAGAGATGCTGTTATTTATGTTATTAATCCGGAATCGT 120
480 GGGAACTGTTTATTTTGGGAGAGATGCTGTTATTTATGTTATTAATCCGGAATCGT 539
QY 121 CACTGAGAAAGTTAAAGAAATTCGAGATCGAATCAACGTAAGACAG 167
|||||
Db 540 CACTGAGAAAGTTAAAGAAATTCGAGATCGAATCAACGTAAGACAG 586

RESULT 12
BQ365363 572 bp mRNA linear EST 21-MAY-2002
LOCUS BR2-GN0027-040900-001-e09 GN0027 Homo sapiens cDNA, mRNA sequence.

DEFINITION BQ365363
ACCESSION BQ365363
VERSION BQ365363.1 GI:21040875
KEYWORDS EST.

SOURCE human.
ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 572)
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

AUTHORS

Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
Goldman,G.H., Carvalho,A.F., Matsukuma,A., Bala,G.S., Simpson,D.H.,
Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
Simpson,A.J.

Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags

JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)

MEDLINE

COMMENT

Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001

Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?l1=MR2&l2=MR2-GN0027-
040900-001-e09&l3=2000-09-04&l4=1)

Seq primer: puc 18 forward
High quality sequence start: 59
High quality sequence stop: 566.

Location/Qualifiers
1. 572

FEATURES

source

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="GN0027"
/dev_stage="Adult"
/note="Organ: placenta.normal; Vector: puc18; Site_1: SmaI
; Site_2: SmaI; A mini-library was made by cloning
products derived from ORESTES PCR (U.S. Letters Patent
application No. 196,716 - Ludwig Institute for Cancer
Research) profiles into the puc 18 vector. Reverse
transcription of tissue mRNA and cDNA amplification were
performed under low stringency conditions."

BASE COUNT

157 a 145 c 140 g 130 t

Query Match 81.2%; Score 163.2; DB 14; Length 572;
Best Local Similarity 94.6%; Pred. No. 1.1e-39;
Matches 191; Conservative 0; Mismatches 8; Indels 3; Gaps 2;

QY 1 CTTAGCAGCAGTACTCTTCAAAATCGAAGAGCTTTAG-ACCTGCTAACCGCTGAAGAG 59
|||||

Db 77 CTTAGCAGCAGTACTCTTCAAAATCGAAGAGCTTTAG-ACCTGCTAACCGCTGAAGAG 136
GGGAACTGTTTATTTTGGGAGAGATGCTGTTATTTATGTTATTAATCCGGAATCG 194

QY 60 GGGAACTGTTTATTTTGGGAGAGATGCTGTTATTTATGTTATTAATCCGGAATCG 119
|||||

Db 137 GGGAACTGTTTATTTTGGGAGAGATGCTGTTATTTATGTTATTAATCCGGAATCG 194
TTTGGGAGAGATGCTGTTATTTATGTTATTAATCCGGAATCG 194

QY 120 TCACTGAGAAAGTTAAAGAAATTCGAGATCGAATCAACGTAAGACAG 179
|||||

Db 195 TCACTGAGAAAGTTAAAGAAATTCGAGATCGAATCAACGTAAGACAG 254
TCACTGAGAAAGTTAAAGAAATTCGAGATCGAATCAACGTAAGACAG 254

QY 180 AACTGGACCCCTGGGCGCTCT 201
|||||

Db 255 ACACGGACCCGGGGCGCTCT 276

RESULT 13
BG004247 302 bp mRNA linear EST 24-JAN-2001
LOCUS RC6-GN0070-301100-014-F12 GN0070 Homo sapiens cDNA, mRNA sequence.

DEFINITION BG004247
ACCESSION BG004247
VERSION BG004247.1 GI:12445227
KEYWORDS EST.

SOURCE human.
ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 302)
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

AUTHORS

Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
Goldman,G.H., Carvalho,A.F., Matsukuma,A., Bala,G.S., Simpson,D.H.,
Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
Simpson,A.J.

Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags

JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)

MEDLINE

COMMENT

Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001

Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?l1=RC6&l2=RC6-GN0070-
301100-014-F12&l3=2000-11-30&l4=1)

Seq primer: puc 18 forward
High quality sequence start: 8
High quality sequence stop: 301.

Location/Qualifiers
1. 302

FEATURES

source

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="GN0070"
/dev_stage="Adult"
/note="Organ: placenta.normal; Vector: puc18; Site_1: SmaI
; Site_2: SmaI; A mini-library was made by cloning
products derived from ORESTES PCR (U.S. Letters Patent
application No. 196,716 - Ludwig Institute for Cancer
Research) profiles into the puc 18 vector. Reverse
transcription of tissue mRNA and cDNA amplification were
performed under low stringency conditions."

BASE COUNT

94 a 69 c 72 g 67 t

Query Match 81.0%; Score 162.8; DB 12; Length 302;
Best Local Similarity 96.0%; Pred. No. 1.3e-39;
Matches 167; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY	1	CGTGCAGCATATCCTCTCAAAATCGAAGCGTTTGTACTGCTAACCCCTAAAGAG	60
Db	129	CGTGCAGCATAGACCCCTTCAAAAGCGAAGCTTTAGACTCTCAACCGCTAAAGAG	188
QY	61	GGGAACCGTTATTTTATTTTAGGGGAAGATCGTGTATATGTTAAATACCGGAATCGT	120
Db	189	GGGAACCTGTTATTTTATTTTAGGGGAAGATCGTGTATATGTTAAATACCGGAATCGT	248
QY	121	CACGTAGAGAAAGTTAAAGAAATTCGAGATCGAATACAACTGAGACAGAGAGCT	174
Db	249	CACGTAGAGAAATTTAAAGAAATCGAGATCGAATACAACTGAGGACAGAGAGCT	302

OY	61	GGGAACCGCTTATTTTAGGGGAGAAATGCTGTATTATGTAACAATCCGAATCGT	120
Db	450	GGGAATCTGTTATTATTTTAGTGGAAGAATGCTGTATTAGCTTATCAATCTGGAAATCGT	391
OY	121	CACGAGAAACTTAAAGAAATTCAGATGCATAATCAACGTAGACAGAGAGCTTGCAAA	180
Db	390	CACCGAANAATTTNAGAAATTCAGATGCCATATACAACTATGACAGAGAGATTTCANAA	331
OY	181	CAC TGAC CCG TGG GCG CTCT 201	
Db	330	CACCAGACCTTGGGACCTCTCT 310	

RESULT 14	
LOCUS	A0725469/c
DEFINITION	A0725469 608 bp DNA linear GSS 14-JUL-1999
ACCESSION	HS-5403_A1-E02-T7A RPCT-11 Human Male BAC Library Homo sapiens
VERSION	A0725469
KEYWORDS	genomic clone PTA-979 Col-3 Row-I, DNA sequence.
SOURCE	A0725469.1 GI:5485138
ORGANISM	GSS.
	human.
	Homo sapiens

RESULT 15	
BG572445	
LOCUS	790 bp mRNA linear EST 10-APR-2001
DEFINITION	602592490P1 NIH_MGC_79 Homo sapiens cDNA clone IMAGE:4720846 5',
	mRNA sequence.
ACCESSION	BC572445
VERSION	BC572445
KEYWORDS	EST.
SOURCE	human.
ORGANISM	Homo sapiens

REFERENCE.
AUTHORS
1 (bases 1 to 608)
Maharaj,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T.,

REFERENCE
1 (bases 1 to 790)
NIH-MGC <http://mgc.ncl.nih.gov/>.

TITLE	Sequence-tagged connectors: A sequence approach to mapping and scanning the human genome
JOURNAL	Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)
MEDLINE	00306500

COMMENT
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: CLONTECH Laboratories, Inc.
Dr. Robert Strausberg, Clontech Laboratories, Inc.

COMMENT
Contact: Manairas GS, Wallace JC, Hood L
High Throughput Sequencing Center
University of Washington
401 Queen Anne Avenue North, Seattle, WA 98109, USA
Tel: (206) 616-3618
Fax: (206) 616-3887
Email: jwallace@u.washington.edu
Clones are derived from the human BAC library RPCI-11. For BAC library availability, please contact Pieter de Jong (pieter@ejonj.med.buffalo.edu). Clones may be purchased from BACPAC Resources (http://BACPAC.med.buffalo.edu/ordering_bac.htm) or from Research Genetics (info@resgen.com). BAC end Web Server: <http://www.hnsc.washington.edu>
plate: 979 row: I column: 3

```

FEATURES
source
    cdna.library.arrayed.by: The I.M.A.G.E. Consortium (LLNL)
    dna.sequencing.by: Incyte Genomics, Inc.
    clone.distribution: MGC clone distribution information can be
    found through the I.M.A.G.E. Consortium/LLNL at:
    http://image.llnl.gov
    plate: LHC1576 row: C column: 23
    High quality sequence stop: 624.
    Location/Qualifiers
        1..790
        /organism="Homo sapiens"
        /db_xref="taxon:9606"
        /clone="IMAGE:4720846"
        /clone_1lb="NIH_MGC_79"
        /lab_host="DH10B (T1 phage-resistant)"

```

```
FEATURES      Location/Qualifiers
source        1 . 608
              /format=sm="Homo sapiens"
```

adaptor sequence: 5'-CACGCCCATTTATGCC-3' and 3' adaptor sequence: 5'-ATCTGAGGCCGAGGCCGACATG-3' (where B = A or C and N = A, C or T).

BASE COUNT
ORIGIN
165 a 114 c 145 g 173 t 11 others

	Library."
BASE COUNT	210 a 202 c 164 g 214 t
ORIGIN	
Query Match	75.88; Score 152.4; DB 12; Length 790;
Best Local Similarity	95.0%; Pred. No. 2.5e-36;
Matches 190; Conservative	0; Mismatches 6; Indels 4; Gaps

Query Match	78.5%	Score 157.8	DB 17	length 608
Best Local Similarity	89.1%	Pred. No. 5.2e-38		
Matches 179	Conservative 0	Mismatches 21	Indels 1	Gaps 1

```
Oy      1 CCTAGCAGCAGTAGTCCTTCA-AAATCGAAGACGTTTGACTTGCCTAAACCCTGAAGAAG 59  
        |||||  
Db     381 CCTAGCAGCAGTAGTCCTTCACAAATCGAAGACGTTTGACTTGCCTAACCCGTGAAGAAG 440
```

```

Qy      1  CCTAGCAGCAGTAGTCCTCAAAATGGAAGACTTTAGCTTCTAACCGCTGAAGAGG  60
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      509  CTTAGCAGCAGTAGTCCTTC-AAATCGAAGAGCTTTAGACTTCTTAACTCGAGAAAGAGA  451

```

Qy 60 GGGGAACCTGTTATTTTAAAGGGAAGATGCTGTATTATGTTATCAATCCGGAATCG 119
|||||
Db 441 GGGGAACCTGTTATTTTAAAGGGAAGATGCTGTATTATGTTATCAATCCGGAATCG 500

OY 120 TCACGTGAAAGTTA--AGAAATTCGAGATCG--AATACACGCTAGACAGAGAGCTTC 176
 |||||
 Db 501 TCACGTGAAAGTTACAGAAATTCGAGATCGCATATACACGTAGACAGAGAGCTTC 560
 OY 177 GAAACACTGGACCCCTGGGGC 196
 |||||
 Db 561 GAAACACTGGACCCCTGGGGC 580

Search completed: May 2, 2003, 16:21:36
 Job time : 1116 secs

GenCore version 5.1.5
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 2, 2003, 14:52:21 ; Search time 39.333 Seconds
(without alignments)
1567.170 Million cell updates/sec

Title: US-09-719-554-3_COPY_9000_9200

Perfect score: 201
Sequence: 1 cctagcagcagtagtcctc.....actgaccctgggacctcct 201

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapect 1.0

Searched: 441362 seqs, 15338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/2/ina/5A.COMB.seq:*
2: /cgn2_6/ptodata/2/ina/5B.COMB.seq:*
3: /cgn2_6/ptodata/2/ina/6A.COMB.seq:*
4: /cgn2_6/ptodata/2/ina/6B.COMB.seq:*
5: /cgn2_6/ptodata/2/ina/PCBUS.COMB.seq:*
6: /cgn2_6/ptodata/2/ina/Backfile1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	201	100.0	2946	US-09-175-928-3	Sequence 3, Appli
2	72.6	36.1	2518	US-09-011-745-1	Sequence 1, Appli
3	72.6	36.1	5865	US-09-011-745-8	Sequence 8, Appli
4	60.6	30.1	8060	US-08-766-528-1	Sequence 1, Appli
5	59	29.4	1704	US-08-007-282B-1	Sequence 1, Appli
6	59	29.4	8132	US-08-766-528-3	Sequence 1, Appli
7	37	28.4	10970	US-08-716-351A-5	Sequence 3, Appli
8	56.6	28.2	1989	US-09-232-278A-5	Sequence 5, Appli
9	56.6	28.2	2499	US-08-105-483-310	Sequence 310, App
10	56.6	28.2	2499	US-08-709-209-310	Sequence 310, App
11	56.6	28.2	2499	US-08-458-101-310	Sequence 310, App
12	54	26.9	1722	US-08-691-563C-58	Sequence 58, Appli
13	54	26.9	1859	US-08-691-563C-46	Sequence 46, Appli
14	51	25.4	8655	US-09-075-272-1	Sequence 1, Appli
15	50.8	25.3	1979	US-08-552-369-19	Sequence 19, Appli
16	49.4	24.6	1493	US-09-376-781-25	Sequence 24, Appli
17	49.4	24.6	2000	US-09-376-781-25	Sequence 25, Appli
18	49.4	24.6	2000	US-09-376-781-30	Sequence 30, Appli
19	49.4	24.6	2462	US-09-111-085-1	Sequence 1, Appli
20	49.4	24.6	7333	US-08-766-528-2	Sequence 2, Appli
21	48.2	24.0	464	US-09-120-653D-6	Sequence 6, Appli
22	48.2	24.0	3910	US-09-120-653D-1	Sequence 1, Appli
23	47.8	23.8	2002	US-09-315-127-7	Sequence 7, Appli
24	47.8	23.8	3925	US-09-011-745-9	Sequence 9, Appli
25	47.8	23.8	5109	US-08-073-836-1	Sequence 1, Appli
26	47.8	23.8	5176	US-08-073-836-2	Sequence 2, Appli
27	47.8	23.8	6061	US-09-011-745-6	Sequence 6, Appli

28	47.8	23.8	6795	US-09-380-190A-22	Sequence 22, Appli
29	47.8	23.8	8332	US-08-850-961-1	Sequence 1, Appli
30	47.8	23.8	8332	US-09-479-776-1	Sequence 1, Appli
31	47.8	23.8	8332	US-09-309-572-11	Sequence 11, Appli
32	47.8	23.8	8332	US-09-315-127-1	Sequence 1, Appli
33	47.8	23.8	8332	US-09-265-013-1	Sequence 1, Appli
34	47.8	23.8	9093	US-09-380-190A-23	Sequence 23, Appli
35	47.4	23.6	6312	US-09-011-745-7	Sequence 7, Appli
36	47.4	23.6	8088	US-09-315-127-4	Sequence 4, Appli
37	47.4	23.6	8535	US-08-716-351A-1	Sequence 1, Appli
38	46.2	23.0	1911	US-08-258-420-8	Sequence 8, Appli
39	46.2	23.0	1965	US-08-258-420-9	Sequence 9, Appli
40	46.2	23.0	2001	US-08-850-961-13	Sequence 13, Appli
41	46.2	23.0	2001	US-09-479-776-13	Sequence 13, Appli
42	46.2	23.0	2001	US-09-315-127-10	Sequence 10, Appli
43	46.2	23.0	6028	US-09-011-745-5	Sequence 5, Appli
44	46.2	23.0	8202	US-08-258-420-13	Sequence 13, Appli
45	45.4	22.6	80246	US-09-078-294-4	Sequence 4, Appli

ALIGNMENTS

```
RESULT 1
US-09-175-928-3
; Sequence 3, Application US/09175928A
; Patent No. 6312921
; GENERAL INFORMATION:
; APPLICANT: Jacobs, Kenneth
; APPLICANT: McCoy, John M.
; APPLICANT: Lavalie, Edward R.
; APPLICANT: Collins-Racle, Lisa A.
; APPLICANT: Evans, Cheryl
; APPLICANT: Werberg, David
; APPLICANT: Treacy, Maurice
; APPLICANT: Mi, Sha
; APPLICANT: Genetics Institute, Inc.
; TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES ENCODING THEM
; FILE REFERENCE: 6006B AJ172A
; CURRENT APPLICATION NUMBER: US/09/175,928A
; CURRENT FILING DATE: 1998-10-20
; NUMBER OF SEQ ID NOS: 62
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 2946
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-175-928-3
Query Match 100.0%; Score 201; DB 4; Length 2946;
Best Local Similarity 100.0%; Pred. No. 3,4e-54;
Matches 201; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CCTACAGCAGTAGCTCTTCAAAATCGAAGACCTTAGACTGCTAACCGCTGAAGAGG 60
DB 2049 CCTAGAGAGTAGTACCTTCAAAATCGAAGACCTTAGACTGCTAACCGCTGAAGAGG 2108
QY 61 GGGAACTGTTATTTTATTTAGGGGAAGATGCTTATTTATTTATTCATCCGAATCGT 120
DB 2109 GGGAACTGTTATTTTATTTAGGGGAAGATGCTTATTTATTTATTCATCCGAATCGT 2168
QY 121 CACTGAGAAAGTTAAGAAATCGAGATGATACAGCTAGAGGAGAGAGCTTGGANA 180
DB 2169 CACTGAGAAAGTTAAGAAATCGAGATGATACAGCTAGAGGAGAGAGCTTGGANA 2228
QY 181 CACTGAGACCTGGGCGCTCT 201
DB 2229 CACTGAGACCTGGGCGCTCT 2249
RESULT 2
US-09-011-745-1
; Sequence 1, Application US/09011745
```

```
Patent No. 6165715
GENERAL INFORMATION:
APPLICANT: COLLINS, Mary KL
APPLICANT: Weiss, Robin A
APPLICANT: Takeuchi, Yasuhiro
APPLICANT: Cosset, Francois-Loic
TITLE OF INVENTION: Expression systems
FILE REFERENCE: 09/011,745
CURRENT APPLICATION NUMBER: US/09/011,745
CURRENT FILING DATE: 1998-06-22
EARLIER APPLICATION NUMBER: PCT/GB96/02061
EARLIER FILING DATE: 1996-08-23
EARLIER APPLICATION NUMBER: GB9517263.1
EARLIER FILING DATE: 1995-08-23
NUMBER OF SEQ ID NOS: 29
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 1
LENGTH: 2518
TYPE: DNA
ORGANISM: RD114
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1)
OTHER INFORMATION: n is any nucleotide
US-09-011-745-1
```

```
Query Match          36.1%; Score 72.6; DB 4; Length 2518;
Best Local Similarity 60.3%; Pred. No. 9.3e-14;
Matches 120; Conservative 0; Mismatches 79; Indels 0; Gaps 0;
```

```
QY 3 TACACGAGTAGTCCTTCAAAATCGAAGAGCTTTAGACTTGTACCGCTGGAAGAGGG 62
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1913 TACCTGAGTAGTCTCCAAAATAGGAGGAGCTGACCTACTACGCGAGAAACAAGAG 1972

QY 63 GAACCTGTTATTTTAGGGGGAAGATGCTTATTATGTTATTCATCCGGAATCGTCA 122
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1973 GAATTTGTTAGCTTACAGAAAATGCTGTTTATGCTACAGTCAGGAATGTGA 2032.

QY 123 CTGAGAAAGTTAAGAAATTCGAGATCGAATACAGCTAGAGCAGAGAGCTTCGAACA 182
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 2033 GAAACAAAATTAAGAACCCCTACAGAAAGAAATTACAAAACGACGAGAAAGCTGGCAACA 2092

QY 183 CTGACCCCTGGGGCTCTCT 201
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 2093 ACCCTCTCTGGACCGGCT 2111
```

RESULT 3

```
US-09-011-745-8
Sequence 8, Application US/09011745
Patent No. 6165715
GENERAL INFORMATION:
APPLICANT: COLLINS, Mary KL
APPLICANT: Weiss, Robin A
APPLICANT: Takeuchi, Yasuhiro
APPLICANT: Cosset, Francois-Loic
TITLE OF INVENTION: Expression systems
FILE REFERENCE: 09/011,745
CURRENT APPLICATION NUMBER: US/09/011,745
CURRENT FILING DATE: 1998-06-22
PCT/GB96/02061
EARLIER APPLICATION NUMBER: PCT/GB96/02061
EARLIER FILING DATE: 1996-08-23
EARLIER APPLICATION NUMBER: GB9517263.1
EARLIER FILING DATE: 1995-08-23
NUMBER OF SEQ ID NOS: 29
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 8
LENGTH: 5865
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Portion of
```

```
FEATURE:
NAME/KEY: misc_feature
LOCATION: (3611)
OTHER INFORMATION: n is any nucleotide
FEATURE:
NAME/KEY: misc_feature
LOCATION: (3612)
OTHER INFORMATION: n is any nucleotide
FEATURE:
NAME/KEY: misc_feature
LOCATION: (3613)
OTHER INFORMATION: n is any nucleotide
FEATURE:
NAME/KEY: misc_feature
LOCATION: (3614)
OTHER INFORMATION: n is any nucleotide
FEATURE:
NAME/KEY: misc_feature
LOCATION: (3799)
OTHER INFORMATION: n is any nucleotide
FEATURE:
NAME/KEY: misc_feature
LOCATION: (3800)
OTHER INFORMATION: n is any nucleotide
FEATURE:
NAME/KEY: misc_feature
LOCATION: (3801)
OTHER INFORMATION: n is any nucleotide
FEATURE:
NAME/KEY: misc_feature
LOCATION: (3802)
OTHER INFORMATION: n is any nucleotide
US-09-011-745-8
```

```
Query Match          36.1%; Score 72.6; DB 4; Length 5865;
Best Local Similarity 60.3%; Pred. No. 1.2e-13;
Matches 120; Conservative 0; Mismatches 79; Indels 0; Gaps 0;
```

```
QY 3 TACACGAGTAGTCCTTCAAAATCGAAGAGCTTTAGACTTGTACCGCTGGAAGAGGG 62
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 2771 TACCTGAGTAGTCTCCAAAATAGGAGGAGCTGACCTACTACGCGAGAAACAAGAG 2830

QY 63 GAACCTGTTATTTTAGGGGGAAGATGCTTATTATGTTATTCATCCGGAATCGTCA 122
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 2831 GAATTTGTTAGCTTACAGAAAATGCTGTTTATGCTACAGTCAGGAATGTGA 2890

QY 123 CTGAGAAAGTTAAGAAATTCGAGATCGAATACAGCTAGAGCAGAGAGCTTCGAACA 182
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 2891 GAAACAAAATTAAGAACCCCTACAGAAAGAAATTACAAAACGACGAGAAAGCTGGCAACA 2950

QY 183 CTGACCCCTGGGGCTCTCT 201
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 2951 ACCCTCTCTGGACCGGCT 2969
```

RESULT 4

```
US-08-766-528-1
Sequence 1, Application US/08766528
Patent No. 6190861
GENERAL INFORMATION:
APPLICANT: Jay A. Fishman
TITLE OF INVENTION: MOLECULAR SEQUENCE OF SWINE RETROVIRUS
TITLE OF INVENTION: AND METHODS OF USE
NUMBER OF SEQUENCES: 74
CORRESPONDENCE ADDRESS:
ADDRESSEE: LAHIVE & COCKFIELD, LLP
STREET: 60 State Street
CITY: Boston
STATE: Massachusetts
COUNTRY: USA
ZIP: 02109-1875
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
```


TOPOLOGY: linear
MOLECULE TYPE: CDNA
US-08-766-528-3

Query Match
Best Local Similarity 62.6%; Pred. No. 2.6e-09;
Matches 92; Conservative 0; Mismatches 55; Indels 0; Gaps 0;

29.4%; Score 59; DB 4; Length 8132;

QY 1 CCTAGCAGCAGTACTCTTCAAAATCGAAGAGCTTTAGACTTGTACCCGCTGAAGAGG 60
DB 7134 CTATCTGAGTAGTCTACAGAAATAGAGAGGGTTAGATTATTTCTAAAGAGG 7193
QY 61 GCGAACCGTTATTTTATAGGGAAGATGCTTTATATGTTAATCATCCGGAATCGT 120
DB 7194 AGGATTAGTGTACCTTGAAGAGAGGATGCTTTTATGATCATTCAGGGGCGCAT 7253
QY 121 CACTGAGAAAGTTAAGAAATTCGAGA 147
DB 7254 CAGAGACTCCATGAACAAGCTTAGAGA 7280

RESULT 7
US-08-716-351A-5
Sequence 5, Application US/08716351A
Patent No. 6033905
GENERAL INFORMATION:
APPLICANT:
TITLE OF INVENTION: Gibbon Ape Leukemia Virus-Based
TITLE OF INVENTION: Retroviral Vectors
NUMBER OF SEQUENCES: 5
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/716.351A
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US94/03784
FILING DATE: 06-APR-1994
ATTORNEY/AGENT INFORMATION:
NAME: Bastian, Kevin L.
REGISTRATION NUMBER: 34,774
REFERENCE/DOCKET NUMBER: 15280-128-1PC
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 543-9600
TELEFAX: (415) 543-5043
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 10970 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: misc_feature
LOCATION: 1..10970
OTHER INFORMATION: /standard_name="p537 retroviral"
OTHER INFORMATION: vector"
US-08-716-351A-5

Query Match
Best Local Similarity 56.0%; Pred. No. 1.2e-08;
Matches 108; Conservative 0; Mismatches 85; Indels 0; Gaps 0;

28.4%; Score 57; DB 3; Length 10970;

QY 1 CCTAGCAGCAGTACTCTTCAAAATCGAAGAGCTTTAGACTTGTACCCGCTGAAGAGG 60
DB 6894 CCTATCTGAGTAGTCTACAGAAATAGAGAGCGCTTGACTTACTATCTTAAAGAGG 6953
QY 61 GCGAACCGTTATTTTATAGGGAAGATGCTTTATATGTTAATCATCCGGAATCGT 120

DB 6954 AGGCTCTGCGCGCCCTTAAAGAGAGTGTGTTTTATGTAGACCAGCTCAGGTGCAGT 7013
QY 121 CACTGAGAAAGTTAAGAAATTCGAGATCGAATCAACAGTAGACAGAGAGCTTCGAAA 180
DB 7014 ACGAGACTCCATGAAAAACTTTAAAGAAAGACTAGATTAAGACAGTAGAGCCCGCAGAA 7073
QY 181 CACTGGAGCCCTGG 193
DB 7074 AAACCAAACTCG 7086

RESULT 8
US-09-232-278A-5
Sequence 5, Application US/09232278A
Patent No. 6348196
GENERAL INFORMATION:
APPLICANT: AUDONNET et al.
TITLE OF INVENTION: FELINE POLYNUCLEOTIDE VACCINE FORMULA
FILE REFERENCE: 454313-2220
CURRENT APPLICATION NUMBER: US/09/232.278A
CURRENT FILING DATE: 1999-01-15
NUMBER OF SEQ ID NOS: 30
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 5
LENGTH: 1989
TYPE: DNA
ORGANISM: Feline leukemia virus
US-09-232-278A-5

Query Match
Best Local Similarity 57.7%; Pred. No. 9.5e-09;
Matches 101; Conservative 0; Mismatches 74; Indels 0; Gaps 0;

28.2%; Score 56.6; DB 4; Length 1989;

QY 1 CCTAGCAGCAGTACTCTTCAAAATCGAAGAGCTTTAGACTTGTACCCGCTGAAGAGG 60
DB 1584 CCTCTCGAGTACTCTTCAAAATAGAGAGCGGCTGATATCTGTTTCAAAAAGG 1643
QY 61 GCGAACCGTTATTTTATAGGGAAGATGCTTTATATGTTAATCATCCGGAATCGT 120
DB 1644 AGGCTCTGTGCGCCCTTAAAGAGAGATGCTGTTATGACAGATACACCGGACTCGT 1703
QY 121 CACTGAGAAAGTTAAGAAATTCGAGATCGAATACAGAGAGCAGAGAGCTT 175
DB 1704 CAGAGACAATATGCTTAATTAAGAGAAAGACTGAACAGCAGACACACTGTTT 1758

RESULT 9
US-08-105-483-310
Sequence 310, Application US/08105483
Patent No. 5494807
GENERAL INFORMATION:
APPLICANT: Paolletti, Enzo
TITLE OF INVENTION: GENETICALLY ENGINEERED VACCINE
TITLE OF INVENTION: STRAIN
NUMBER OF SEQUENCES: 462
CORRESPONDENCE ADDRESS:
ADDRESSEE: Curtis, Morris & Safford
ADDRESSEE: c/o William S. Frommer
STREET: 530 Fifth Avenue
CITY: New York
STATE: NY
COUNTRY: USA
ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/105.483
FILING DATE: 12-AUG-1993
CLASSIFICATION: 424
PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/847,951
FILING DATE: 06-MAR-1992
ATTORNEY/AGENT INFORMATION:
NAME: Frommer, William S.
REGISTRATION NUMBER: 25,506
REFERENCE/DOCKET NUMBER: 454310-2400
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 840-3333
TELEFAX: (212) 840-0712
INFORMATION FOR SEQ ID NO: 310:
SEQUENCE CHARACTERISTICS:
LENGTH: 2499 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-105-483-310

Query Match 28.2%; Score 56.6; DB 1; Length 2499;
Best Local Similarity 57.7%; Pred. No. 1e-08;
Matches 101; Conservative 0; Mismatches 74; Indels 0; Gaps 0;

QY 1 CCTAGACAGTAGTCCTCAAAATCGAAGAGCTTAGACTGCTAACCGCTGAAGAG 60
DB 1608 CCTCTCCGAGGTAGCTTACAAATAGACGGGCTTAGATATTCTTCTTACAAAGG 1667

QY 61 GGAACCTGTTATTTTATGAGGAGATGCTGTATTTATGTAATCAATCCGAAATCGT 120
DB 1668 AGGGCTCTGTGCCCTTAAAGGAGAAATGCTTCTTATGAGATCACCACCGACTCGT 1727

QY 121 CACTGAGAAAGTTAAAGAAATTCGAGATCGAATACACAGCTAGAGCAGAGAGCTT 175
DB 1728 CAGAGACAATATGCTAATTAAGAGAAAGACTGAACAGCGACACCAACTGTTT 1782

RESULT 10
US-08-709-209-310
Sequence 310, Application US/08709209
Patent No. 5762938
GENERAL INFORMATION:
APPLICANT: Paoletti, Enzo
TITLE OF INVENTION: GENETICALLY ENGINEERED VACCINE
TITLE OF INVENTION: STRAIN
NUMBER OF SEQUENCES: 462
CORRESPONDENCE ADDRESS:
ADDRESSEE: Curtis, Morris & Safford
STREET: 530 Fifth Avenue
CITY: New York
STATE: NY
COUNTRY: USA
ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/709,209
FILING DATE: 21-AUG-1996
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/105,483
FILING DATE: 12-AUG-1993
APPLICATION NUMBER: US 07/847,951
FILING DATE: 06-MAR-1992
ATTORNEY/AGENT INFORMATION:
NAME: Frommer, William S.
REGISTRATION NUMBER: 25,506
REFERENCE/DOCKET NUMBER: 454310-2400
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 840-3333
TELEFAX: (212) 840-0712
INFORMATION FOR SEQ ID NO: 310:

SEQUENCE CHARACTERISTICS:
LENGTH: 2499 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-709-209-310

Query Match 28.2%; Score 56.6; DB 1; Length 2499;
Best Local Similarity 57.7%; Pred. No. 1e-08;
Matches 101; Conservative 0; Mismatches 74; Indels 0; Gaps 0;

QY 1 CCTAGACAGTAGTCCTCAAAATCGAAGAGCTTAGACTGCTAACCGCTGAAGAG 60
DB 1608 CCTCTCCGAGGTAGCTTACAAATAGACGGGCTTAGATATTCTTCTTACAAAGG 1667

QY 61 GGAACCTGTTATTTTATGAGGAGATGCTGTATTTATGTAATCAATCCGAAATCGT 120
DB 1668 AGGGCTCTGTGCCCTTAAAGGAGAAATGCTTCTTATGAGATCACCACCGACTCGT 1727

QY 121 CACTGAGAAAGTTAAAGAAATTCGAGATCGAATACACAGCTAGAGCAGAGAGCTT 175
DB 1728 CAGAGACAATATGCTAATTAAGAGAAAGACTGAACAGCGACACCAACTGTTT 1782

RESULT 11
US-08-458-101-310
Sequence 310, Application US/08458101
Patent No. 5766599
GENERAL INFORMATION:
APPLICANT: Paoletti, Enzo
TITLE OF INVENTION: GENETICALLY ENGINEERED VACCINE
TITLE OF INVENTION: STRAIN
NUMBER OF SEQUENCES: 467
CORRESPONDENCE ADDRESS:
ADDRESSEE: Curtis, Morris & Safford
STREET: 530 Fifth Avenue
CITY: New York
STATE: NY
COUNTRY: USA
ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/458,101
FILING DATE: 01-JUN-1995
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Frommer, William S.
REGISTRATION NUMBER: 25,506
REFERENCE/DOCKET NUMBER: 454310-2740
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 840-3333
TELEFAX: (212) 840-0712
INFORMATION FOR SEQ ID NO: 310:
SEQUENCE CHARACTERISTICS:
LENGTH: 2499 base pairs
TYPE: nucleic acid

STRANDEDNESS: single
TOPOLOGY: linear
US-08-458-101-310

Query Match 28.2%; Score 56.6; DB 1; Length 2499;
Best Local Similarity 57.7%; Pred. No. 1e-08;
Matches 101; Conservative 0; Mismatches 74; Indels 0; Gaps 0;

QY 1 CCTACGACGATGCTCTCAAAATCGAAGACCTTAGACTGCTACCGCGTAAGAGG 60
DB 1608 CCTCTCCGAGTAGTCTTACAAAATAGACGGGCGCTAGATATCTGTTCTTACAAAAGG 1667
QY 61 GGAACCTGTTATTTTGGGGAAGATGCTTATATGTTATTCATCCGGAATCGT 120
DB 1668 AGGCTCTGTGCGCCCTTAAGGAAGATGCTGCTTCTATGCAATCACCGGACTCGT 1727
QY 121 CACTGAGAAAGTTAAAGAAATTCGAGATCGAATACAGCTAGAGGAGACTT 175
DB 1728 CAGAGACAATATGCTAATTAAGAAAGACTGAACGACGACAACAAGTCTT 1782

RESULT 12

US-08-691-563C-58
Sequence 58, Application US/08691563C
Patent No. 6001987

GENERAL INFORMATION:

APPLICANT: Herve PERRON
APPLICANT: Frederic BESEME
APPLICANT: Frederic BEDIN
APPLICANT: Glaucia PARANHOS-BACCALA
APPLICANT: Florence KOMURIAN-PRADEL
APPLICANT: Colette JOLIVET
APPLICANT: Bernard MANDRAND
TITLE OF INVENTION: VIRAL MATERIAL AND NUCLEOTIDE FRAGMENTS
TITLE OF INVENTION: ASSOCIATED WITH MULTIPLE SCLEROSIS, FOR DIAGNOSTIC, PROPHYLACTIC
TITLE OF INVENTION: THERAPEUTIC PURPOSES
NUMBER OF SEQUENCES: 92
CORRESPONDENCE ADDRESSES:
ADDRESS: Oliff & Berridge
STREET: 700 South Washington Street, Suite 300
CITY: Alexandria
STATE: Virginia
COUNTRY: U.S.A.
ZIP: 22314

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/691,563C
FILING DATE: 02-AUG-1996
ATTORNEY/AGENT INFORMATION:
NAME: Berridge, William P.
REGISTRATION NUMBER: 30,024
REFERENCE/DOCKET NUMBER: WPB 38588
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-836-6400
TELEFAX: 703-836-2787

INFORMATION FOR SEQ ID NO: 58:

SEQUENCE CHARACTERISTICS:
LENGTH: 1722 base pairs
TYPE: nucleotide
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
US-08-691-563C-58

Query Match 26.9%; Score 54; DB 3; Length 1722;
Best Local Similarity 66.1%; Pred. No. 6e-08;
Matches 78; Conservative 0; Mismatches 40; Indels 0; Gaps 0;

QY 3 TAGACGACGATGCTCTCAAAATCGAAGACCTTAGACTGCTAAGCGGTGAAGAGGG 62

DB 530 TGGCAGCATGACTCTCCAAAACCCGAGGCCACACCTCTCAGTGTGAGAAAGAG 589

QY 63 GAACCTGTTATTTTAAAGGAAGATGCTGTTATATGTTAATCAATCCGAATCGT 120
DB 590 GACTCTGCACCTTCTTAAGGGAAGATGTTGTTTACACTACCAAGTGAAGATAGT 647

RESULT 13

US-08-691-563C-46
Sequence 46, Application US/08691563C
Patent No. 6001987

GENERAL INFORMATION:

APPLICANT: Herve PERRON
APPLICANT: Frederic BESEME
APPLICANT: Frederic BEDIN
APPLICANT: Glaucia PARANHOS-BACCALA
APPLICANT: Florence KOMURIAN-PRADEL
APPLICANT: Colette JOLIVET
APPLICANT: Bernard MANDRAND
TITLE OF INVENTION: VIRAL MATERIAL AND NUCLEOTIDE FRAGMENTS
TITLE OF INVENTION: ASSOCIATED WITH MULTIPLE SCLEROSIS, FOR DIAGNOSTIC, PROPHYLACTIC
TITLE OF INVENTION: THERAPEUTIC PURPOSES
NUMBER OF SEQUENCES: 92
CORRESPONDENCE ADDRESSES:
ADDRESS: Oliff & Berridge
STREET: 700 South Washington Street, Suite 300
CITY: Alexandria
STATE: Virginia
COUNTRY: U.S.A.
ZIP: 22314

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/691,563C
FILING DATE: 02-AUG-1996
ATTORNEY/AGENT INFORMATION:
NAME: Berridge, William P.
REGISTRATION NUMBER: 30,024
REFERENCE/DOCKET NUMBER: WPB 38588
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-836-6400
TELEFAX: 703-836-2787

INFORMATION FOR SEQ ID NO: 46:

SEQUENCE CHARACTERISTICS:
LENGTH: 1859 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
US-08-691-563C-46

Query Match 26.9%; Score 54; DB 3; Length 1859;
Best Local Similarity 66.1%; Pred. No. 6.1e-08;
Matches 78; Conservative 0; Mismatches 40; Indels 0; Gaps 0;

QY 3 TAGACGACGATGCTCTCAAAATCGAAGACCTTAGACTGCTAAGCGGTGAAGAGGG 62
DB 1180 TGGCAGCATGACTCTCCAAAACCGTCAAGGCCCTCAGCTCTCAGTGTGAGAAAGAG 1239

QY 63 GAACCTGTTATTTTAAAGGAAGATGCTGTTATATGTTAATCAATCCGAATCGT 120
DB 1240 GACTCTGCACCTTCTTAAGGGAAGATGTTGTTTACACTACCAAGTGAAGATAGT 1297

RESULT 14

US-09-075-272-1
Sequence 1, Application US/09075272
Patent No. 6136598
GENERAL INFORMATION:

APPLICANT: MILLER, A. DUSTY
APPLICANT: WOLGAST, GREG
APPLICANT: BONHAM, LYNN
TITLE OF INVENTION: MUS DUNNI ENDOGENOUS RETROVIRAL
TITLE OF INVENTION: PACKAGING CELL LINES
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESS:
ADDRESSEE: TOWNSEND and TOWNSEND and CREW LLP
STREET: Two Embarcadero Center, 8th floor
City: San Francisco
STATE: California
COUNTRY: U.S.A.
ZIP: 94111
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/075,272
FILING DATE: 08-MAY-1998
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/046,140
FILING DATE: 09-MAY-1997
ATTORNEY/AGENT INFORMATION:
NAME: POOT, Brian W.
REGISTRATION NUMBER: 32,928
REFERENCE/DOCKET NUMBER: 1458A-003710
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 467-9600
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ. ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 8655 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
US-09-075-272-1

Query Match 25.4%; Score 51; DB 3; Length 8655;
Best Local Similarity 58.1%; Pred. No. 8.6e-07;
Matches 90; Conservative 0; Mismatches 65; Indels 0; Gaps 0;

QY 2 CTAGACAGTAGTCTTCAAAATCGAAGCTTTAGACTTGTACCGCTGAAAGAGG 61
DB 7374 CTGCTGAGGTAGTACTGCAAAATAGGAGGATTTAGATTGTTCTTAAGAAGA 7433

QY 62 GGAACCTGTTATTTTAAAGGGAAGATGCTTATATGTTATTCATTCAGCAATCGTC 121
DB 7434 GGAATTGCTGCTGCTCAAGAGGAATGTTGTTCTATGTCGACCTCGGAGTGATC 7493

QY 122 ACTGAGAACTTAAGAATTCGATCGAATACA 156
DB 7494 AAGATTCTATGGCCAACTTAGAGAACGCCCTAGA 7528

RESULT 15
US-08-552-369-19
; Sequence 19, Application US/08552369
; Patent No. 6241989
; GENERAL INFORMATION:
; APPLICANT: Scott, Fred W.
; APPLICANT: Ngichabe, Christopher K.
; APPLICANT: Hu, Liangbiao
; TITLE OF INVENTION: Recombinant Multivalent Viral Vaccine
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hodgson, Russ, Andrews, Woods & Goodyear
; STREET: 1800 One Met Plaza
; City: Buffalo
; STATE: New York

COUNTRY: United States
ZIP: 14203-2391
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.5 inch, 1.44 Kb storage
COMPUTER: IBM compatible
OPERATING SYSTEM: MS-DOS/ Microsoft Windows 3.1
SOFTWARE: Wordperfect
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/552,369
FILING DATE:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/190,789
FILING DATE: 01/27/1994
APPLICATION NUMBER: 07/726,609
FILING DATE: 07/09/1991
ATTORNEY/AGENT INFORMATION:
NAME: Nelson, M. Bud
REGISTRATION NUMBER: 35,300
REFERENCE/DOCKET NUMBER: 18617.0016
TELECOMMUNICATION INFORMATION:
TELEPHONE: (716) 849-0349
TELEFAX: (716) 849-0349
INFORMATION FOR SEQ. ID NO: 19:
SEQUENCE CHARACTERISTICS:
LENGTH: 1979 nucleotides
TYPE: nucleic acid
STRANDEDNESS: double-stranded
TOPOLOGY: linear
MOLECULE TYPE: DNA
HYPOTHETICAL: Yes
ORIGINAL SOURCE:
ORGANISM: feline leukemia virus
FEATURE:
LOCATION: env gene region
US-08-552-369-19

Query Match 25.3%; Score 50.8; DB 4; Length 1979;
Best Local Similarity 57.6%; Pred. No. 6.4e-07;
Matches 91; Conservative 0; Mismatches 67; Indels 0; Gaps 0;

QY 1 CCTAGCAGTAGTCTTCAAAATCGAAGCTTTAGACTTGTACCGCTGAAAGAGG 60
DB 1574 CCTTTCGAGTAGTCTTACAAACAGCGGCTGATATCTTATCTTACAAAGGG 1633

QY 61 GGAACCTGTTATTTTAAAGGGAAGATGCTTATATGTTATTCATTCAGCAATCGTC 120
DB 1634 AGGCTCTGTGCCCCCATTTGAAGAAGATGTTGCTCTATGCGGATCACACCGGACTCGT 1693

QY 121 CACTGAGAAAGTTAAGAATTCGAGATCGAATACAC 158
DB 1694 CCGAGACATATGCGCAAAATTAGAGAAAGACTAAAC 1731

Search completed: May 2, 2003, 16:24:15
Job time : 53.3333 secs

GenCore version 5.1.5
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 2, 2003, 16:21:41 ; Search time 63.333 Seconds
(without alignments)
3750.092 Million cell updates/sec

Title: US-09-719-554-3_COPY_9000_9200

Perfect score: 201
Sequence: 1 cctagcagcagtagtctctc.....actgacctggggcctctc 201

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 746064 seqs, 590810554 residues

Total number of hits satisfying chosen parameters: 1492128

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published_Applications_NA:*

- 1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
- 2: /cgn2_6/ptodata/1/pubpna/PCr_NEM_PUB.seq:*
- 3: /cgn2_6/ptodata/1/pubpna/US06_NEM_PUB.seq:*
- 4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
- 5: /cgn2_6/ptodata/1/pubpna/US07_NEM_PUB.seq:*
- 6: /cgn2_6/ptodata/1/pubpna/PCrUS_PUBCOMB.seq:*
- 7: /cgn2_6/ptodata/1/pubpna/US08_NEM_PUB.seq:*
- 8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
- 9: /cgn2_6/ptodata/1/pubpna/US09_NEM_PUB.seq:*
- 10: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
- 11: /cgn2_6/ptodata/1/pubpna/US10_NEM_PUB.seq:*
- 12: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*
- 13: /cgn2_6/ptodata/1/pubpna/US60_NEM_PUB.seq:*
- 14: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	201	100.0	2930	10 US-09-902-535-1	Sequence 1, Appl1
2	201	100.0	2946	9 US-10-114-893-134	Sequence 134, App
3	183.4	91.2	1481	7 US-08-979-847-105	Sequence 105, App
4	178.6	88.9	1894	10 US-09-864-761-4444	Sequence 4444, App
5	163.8	81.5	1329	7 US-08-979-847-108	Sequence 108, App
6	127.6	63.5	426	10 US-09-864-761-14030	Sequence 14030, A
7	97.4	48.5	152	10 US-09-864-761-30590	Sequence 30590, A
8	89.4	44.5	520	10 US-09-864-761-7310	Sequence 7310, App
9	84.4	42.0	559	10 US-09-864-761-7501	Sequence 7501, App
10	73.4	36.5	960	9 US-09-764-891-283	Sequence 283, App
11	73.4	36.5	5434	9 US-10-001-835-58	Sequence 58, Appl1
12	71.8	35.7	310	10 US-09-864-761-24000	Sequence 24000, A
13	71.8	35.7	556	10 US-09-864-761-7269	Sequence 7269, App
14	65.8	32.7	1783	10 US-09-729-674-125	Sequence 125, App
15	65.8	32.7	2342	10 US-09-739-254-12	Sequence 12, Appl1
16	65.8	32.7	2342	10 US-09-904-615-12	Sequence 12, Appl1
17	65.8	32.7	5443	9 US-09-764-891-9093	Sequence 9093, App
18	64.6	32.1	143068	10 US-09-967-768A-316	Sequence 316, App
19	58	28.9	8928	9 US-09-764-891-6916	Sequence 6916, App

20	58	28.9	8928	9 US-09-764-891-9705	Sequence 9705, App
21	55	27.4	551	10 US-09-864-761-6686	Sequence 6686, App
22	54	26.9	1722	7 US-08-979-847-54	Sequence 54, Appl1
23	54	26.9	1859	7 US-08-979-847-42	Sequence 42, Appl1
24	49.4	24.6	1493	10 US-09-851-859A-24	Sequence 24, Appl1
25	49.4	24.6	2000	10 US-09-851-859A-25	Sequence 25, Appl1
26	49.4	24.6	2000	10 US-09-851-859A-30	Sequence 30, Appl1
27	48	23.9	792	10 US-09-864-761-21192	Sequence 21192, A
28	47.8	23.8	5130	10 US-09-887-006-9	Sequence 9, Appl1
29	47.8	23.8	8332	10 US-09-006-298-1	Sequence 1, Appl1
30	47.4	23.6	481	9 US-09-918-995-11010	Sequence 11010, A
31	47	23.4	3027	10 US-09-974-298-177	Sequence 177, App
32	45.8	22.8	532	10 US-09-998-598-421	Sequence 421, App
33	45	22.4	505	10 US-09-998-598-384	Sequence 384, App
34	45	22.4	529	10 US-09-920-300A-997	Sequence 997, App
35	45	22.4	529	10 US-10-033-528-997	Sequence 997, App
36	45	22.4	599	10 US-09-998-598-533	Sequence 533, App
37	44.4	22.1	395	10 US-09-864-761-31860	Sequence 31860, A
38	44.4	22.1	579	10 US-09-864-761-15338	Sequence 15338, A
39	44.2	22.0	454	10 US-09-998-598-471	Sequence 471, App
40	43.8	21.8	491	10 US-09-864-761-5518	Sequence 5518, App
41	43.6	21.7	539	10 US-09-919-580-156	Sequence 156, App
42	43.4	21.6	299	10 US-09-998-598-571	Sequence 571, App
43	43.4	21.6	310	10 US-09-998-598-490	Sequence 490, App
44	43.4	21.6	405	10 US-09-998-598-891	Sequence 891, App
45	43.4	21.6	459	10 US-09-998-598-552	Sequence 552, App

ALIGNMENTS

```

RESULT 1
US-09-902-535-1
; Sequence 1, Application US/09902535
; Patent No. US20020102530A1
; GENERAL INFORMATION:
; APPLICANT: Keith, Jr., James C.
; APPLICANT: McCoy, John M.
; TITLE OF INVENTION: Methods and compositions for diagnosing
; TITLE OF INVENTION: and treating preeclampsia and gestational trophoblast
; FILE REFERENCE: GIN-600684
; CURRENT APPLICATION NUMBER: US/09/902,535
; CURRENT FILING DATE: 2001-07-09
; PRIOR APPLICATION NUMBER: 60/216,657
; PRIOR FILING DATE: 2000-07-06
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 2930
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (930)...(2546)
US-09-902-535-1

Query Match      100.0%; Score 201; DB 10; Length 2930;
Best Local Similarity 100.0%; Pred. No. 2.9e-52;
Matches 201; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 CCTACAGCAGTGTCTTCAAAATCGAAGAGCTTTAGACTTGTCAACCGCTGAAGAG 60
DB 2051 CCTACAGCAGTGTCTTCAAAATCGAAGAGCTTTAGACTTGTCAACCGCTGAAGAG 2110
OY 61 GGGACCGTGTATTTTAAAGGAGAGTGTATATGTATATCAATCCGAAATCGT 120
DB 2111 GGGACCGTGTATTTTAAAGGAGAGTGTATATGTATATCAATCCGAAATCGT 2170
OY 121 CACTGAGAAAGTAAAGAAATTCGATTCGATTAACAGCTAGAGAGAGAGCTTCGAAA 180
DB 2171 CACTGAGAAAGTAAAGAAATTCGATTCGATTAACAGCTAGAGAGAGAGCTTCGAAA 2230

```

Oy 181 CACTGACCGTGGGCTCTCT 201
|||
Db 2231 CACTGACCGTGGGCTCTCT 2251

RESULT 2

US-10-114-893-134
; Sequence 134, Application US/01114893
; Publication No. US20020193567A1
; GENERAL INFORMATION:
; APPLICANT: Jacobs, Kenneth
; APPLICANT: McCoy, John M.
; APPLICANT: Lavalie, Edward R.
; APPLICANT: Collins-Racie, Lisa A.
; APPLICANT: Evans, Cheryl
; APPLICANT: Werberg, David
; APPLICANT: Treacy, Maurice
; APPLICANT: Bowman, Michael R.
; APPLICANT: Spaulding, Vikki
; APPLICANT: Carlin-Duckett, McKeough
; APPLICANT: Kelleher, Kerry S.
; APPLICANT: Genetics Institute, Inc.
; TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES ENCODING THEM
; FILE REFERENCE: GI 6000-10A
; CURRENT APPLICATION NUMBER: US/10/114,893
; EARLIER FILING DATE: 2002-04-02
; EARLIER FILING DATE: 1999-10-06
; NUMBER OF SEQ ID NOS: 321
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 134
; LENGTH: 2946
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-114-893-134

Query Match 100.0%; Score 201; DB 9; Length 2946;
Best Local Similarity 100.0%; Pred. No. 2.9e-52;
Matches 201; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 CCTAGCAGCAGTACTCTCTCAAAATCGAAGAGCTTTAGACTTGTACCGCTGAAGAGG 60
|||
Db 2049 CCTAGCAGCAGTACTCTCTCAAAATCGAAGAGCTTTAGACTTGTACCGCTGAAGAGG 2108
Oy 61 GGAACCTGTTATTTTATTTAGGGAAGATGCTTTATTTATGTTATCAATCCGAATCGT 120
|||
Db 2109 GGAACCTGTTATTTTATTTAGGGAAGATGCTTTATTTATGTTATCAATCCGAATCGT 2168
Oy 121 CACTGAGAAAGTTAAAGAAATTCGATTCGATTCGAATTCAGACGACAGAGAGCTTCGAAA 180
|||
Db 2169 CACTGAGAAAGTTAAAGAAATTCGATTCGATTCGAATTCAGACGACAGAGAGCTTCGAAA 2228
Oy 181 CACTGACCGTGGGCTCTCT 201
|||
Db 2229 CACTGACCGTGGGCTCTCT 2249

RESULT 3

US-08-979-847-105
; Sequence 105, Application US/08979847
; Publication No. US20030039664A1
; GENERAL INFORMATION:
; APPLICANT: PERRON, HERVE
; APPLICANT: BESEME, FREDERIC
; APPLICANT: BEDIN, FREDERIC
; APPLICANT: PARANHOS-BACALA, GLAUCIA
; APPLICANT: KOMURIAN-PRADEL, FLORENCE
; APPLICANT: JOLIVET-REYNAUD, COLETTE
; APPLICANT: MANDRAND, BERNARD
; APPLICANT: GARSON, JEREMY
; APPLICANT: TURE, PHILIP
; TITLE OF INVENTION: VIRAL MATERIAL AND NUCLEOTIDE FRAGMENTS

;; TITLE OF INVENTION: ASSOCIATED WITH MULTIPLE SCLEROSIS, FOR DIAGNOSTIC, PROPHY
;; TITLE OF INVENTION: THERAPEUTIC PURPOSES
;; NUMBER OF SEQUENCES: 210
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: OLIVE & BERRIDGE, PLC
;; STREET: P. O. BOX 19928
;; CITY: ALEXANDRIA
;; STATE: VA
;; COUNTRY: USA
;; ZIP: 22320

COMPUTER READABLE FORM:

;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: PatentIn Release #1.0, Version #1.30
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/979,847
;; FILING DATE: 26-NOV-1997
;; CLASSIFICATION: 435
;; ATTORNEY/AGENT INFORMATION:
;; NAME: BERRIDGE, WILLIAM P.
;; REGISTRATION NUMBER: 30,024
;; REFERENCE/DOCKET NUMBER: WPB 39046A
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 703-836-6400
;; TELEFAX: 703-836-2787
;; INFORMATION FOR SEQ ID NO: 105:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 1481 base pairs
;; TYPE: nucleic acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; MOLECULE TYPE: cDNA
US-08-979-847-105

Query Match 91.2%; Score 183.4; DB 7; Length 1481;
Best Local Similarity 94.5%; Pred. No. 6.1e-47;
Matches 190; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

Oy 1 CCTAGCAGCAGTACTCTCTCAAAATCGAAGAGCTTTAGACTTGTACCGCTGAAGAGG 60
|||
Db 1122 CCTAGCAGCAGTACTCTCTCAAAATCGAAGAGCTTTAGACTTGTACCGCTGAAGAGG 1181
Oy 61 GGAACCTGTTATTTTATTTAGGGAAGATGCTTTATTTATGTTATCAATCCGAATCGT 120
|||
Db 1182 GGAACCTGTTATTTTATTTAGGGAAGATGCTTTATTTATGTTATCAATCCGAATCGT 1241
Oy 121 CACTGAGAAAGTTAAAGAAATTCGATTCGATTCGAATTCAGACGACAGAGAGCTTCGAAA 180
|||
Db 1242 CACTGAGAAAGTTAAAGAAATTCGATTCGATTCGAATTCAGACGACAGAGAGCTTCGAAA 1301
Oy 181 CACTGACCGTGGGCTCTCT 201
|||
Db 1302 CACTGACCGTGGGCTCTCT 1322

RESULT 4

US-09-864-761-4444
; Sequence 4444, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FO
; FILE REFERENCE: AeoMica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456

;; PRIOR FILING DATE: 2000-05-26
;; PRIOR APPLICATION NUMBER: US 09/632,366
;; PRIOR FILING DATE: 2000-08-03
;; PRIOR APPLICATION NUMBER: GB 24263.6
;; PRIOR FILING DATE: 2000-10-04
;; PRIOR APPLICATION NUMBER: US 60/236,359
;; PRIOR FILING DATE: 2000-09-27
;; PRIOR APPLICATION NUMBER: PCT/US01/00666
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00667
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00664
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00669
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00665
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00668
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00663
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00662
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00661
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00670
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: US 60/234,687
;; PRIOR FILING DATE: 2000-09-21
;; PRIOR APPLICATION NUMBER: US 09/608,408
;; PRIOR FILING DATE: 2000-06-30
;; PRIOR APPLICATION NUMBER: US 09/774,203
;; PRIOR FILING DATE: 2001-01-29
;; NUMBER OF SEQ ID NOS: 49117
;; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
;; SEQ ID NO 4444
;; LENGTH: 1894
;; TYPE: DNA
;; ORGANISM: Homo sapiens
;; FEATURE:
;; OTHER INFORMATION: MAP TO AC002346.1
;; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 20
;; OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 6
;; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 5.4
;; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 5.9
;; OTHER INFORMATION: EXPRESSED IN PETAL LIVER, SIGNAL = 4
;; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 6.7
;; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 6.2
;; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 6.3
;; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 3.9
;; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 7.2
US-09-864-761-4444

Query Match 88.9%; Score 178.6; DB 10; Length 1894;
Best Local Similarity 93.0%; Pred. No. 2.1e-45;
Matches 187; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

QY 1 CCGAGCAGCAGTCTCTTAATAATGCAAGCTTTAGATTCGTAACCGCTGAAGAGG 60
DB 943 CCGAGCAGCAGTCTCTTAATAATGCAAGCTTTAGATTCGTAACCGCTGAAGAGG 1002
QY 61 GGGAACTGTTATTTTAAAGGAGAGATCTTTATTTATGTAATCAATCCGAATCGT 120
DB 1003 GGGAACTGTTATTTTAAAGGAGAGATCTTTATTTATGTAATCAATCCGAATCAT 1062
QY 121 CACTGAGAAAGTTAAAGAAATTCGAGATCGAATACACGTAAGACAGAGAGCTTCGAAA 180
DB 1063 CACCGAGAAAGTTAAAGAAATTCGAGATCGAATACACGTAAGACAGAGAGCTTCGAAA 1122
QY 181 CACTGAGACCTGGGGCTCTCT 201
DB 1123 CACTGAGACCTGGGGCTCTCT 1143

RESULT 5
US-08-979-847-108

;; Sequence 108, Application US/08979847
;; Publication No. US20030039664A1

;; GENERAL INFORMATION:

;; APPLICANT: PERRON, HERVE
;; APPLICANT: BESEME, FREDERIC

;; APPLICANT: BEDIN, FREDERIC
;; APPLICANT: PARANHOS-BACCALA, GLAUCIA

;; APPLICANT: KOMURIAN-PRADEL, FLORENCE
;; APPLICANT: JOLIVET-REYNAUD, COLETTE

;; APPLICANT: MANDRAND, BERNARD
;; APPLICANT: GARSON, JEREMY

;; APPLICANT: TURE, PHILIP
;; TITLE OF INVENTION: VIRAL MATERIAL AND NUCLEOTIDE FRAGMENTS

;; TITLE OF INVENTION: ASSOCIATED WITH MULTIPLE SCLEROSIS, FOR DIAGNOSTIC, PROPHYL

;; NUMBER OF SEQUENCES: 210
;; CORRESPONDENCE ADDRESS:

;; ADDRESSEE: OLIFF & BERRIDGE, PLC
;; STREET: P.O. BOX 19928

;; CITY: ALEXANDRIA
;; STATE: VA
;; COUNTRY: USA

;; ZIP: 22320
;; COMPUTER READABLE FORM:

;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible

;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: Patent In Release #1.0, Version #1.30

;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/979,847

;; FILING DATE: 26-NOV-1997
;; CLASSIFICATION: 435

;; ATTORNEY/AGENT INFORMATION:
;; NAME: BERRIDGE, WILLIAM P.

;; REGISTRATION NUMBER: 30,024
;; REFERENCE/DOCKET NUMBER: WPB 39046A

;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 703-836-6400

;; TELEFAX: 703-836-2787
;; INFORMATION FOR SEQ ID NO: 108:

;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 1329 base pairs

;; TYPE: nucleic acid
;; STRANDEDNESS: single

;; TOPOLOGY: linear
;; MOLECULE TYPE: cdna
US-08-979-847-108

Query Match 81.5%; Score 163.8; DB 7; Length 1329;
Best Local Similarity 93.4%; Pred. No. 6.9e-41;
Matches 171; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

QY 19 TCAGAAATCGAAGAGCTTACAGCTTCAACCGCGTAAGAGAGGGGGAACCGTTATTTT 78
DB 1 TCAGAAATCGAAGAGCTTACAGCTTCAACCGCGTAAGAGAGGGGGAACCGTTATTTT 60
QY 79 AGGGGAAGATGCTGTTATTTATGTTAATCAATCCGGAATCGTCACTGAGAAAGTTAAGA 138
DB 61 AGGGGAAGATGCTGTTATTTATGTTAATCAATCCGGAATCATTAAGTAAGAAAGTTAAGA 120
QY 139 AATTGAGATCGAATACACGTAAGAGAGAGCTTCGAAACACTGACCTGGGGGCTT 198
DB 121 AATTGAGATCGAATACATTAATGAGAGAGAGAGCTTCGAAACACTGACCTGGGGGCTT 180
QY 199 CCT 201
DB 181 CCT 183

RESULT 6

US-09-864-761-14030
; Sequence 14030, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
; FILE REFERENCE: Aecm1ca-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263,6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 14030
; LENGTH: 426
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AL022067.1
; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 0.95
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.1
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.3
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.2
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 7.1
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.6
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.2
; OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 1.1
US-09-864-761-14030
Query Match 63.5%; Score 127.6; DB 10; Length 426;
Best Local Similarity 80.6%; Pred. No. 6.9e-30;
Matches 187; Conservative 0; Mismatches 14; Indels 31; Gaps 2;
1 CCTACGACGAGTGTCTTCAAAATCGAAGAGCTTTAGACTTGTACCGCTGAAGAGG 60
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

Db 36 CCTACGACGAGTGTCTTCAAAATCGAAGAGCTTTAGACTTGTACCGCTGAAGAGG 95
QY 61 GGAACCTGTTATTTTAGG-----GGAAGATG 90
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 96 GGAAGCTGTTATTTTAGGGAAGATGTTATTTATGTTATTTAGCGGAAGATG 155
QY 91 CTGTTATATGTTATTCATCCGAAATCGTCACTAGAAAGTTAAAGAAATTCGATCG 150
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 156 TGTATTATATGTTATTCATCTGAATGTGCACAGAGAAGTTGAAGAAATTCGATTTG 215
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
QY 151 AATACAGCTAGACGAGAGGACTTC-GAAGACTGGAACCTGGGGCTCCT 201
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 216 AATACAGCTAGACGAGAGGACTTCAAAACACCAAGCCCTGGGGCTCCT 267
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
RESULT 7
US-09-864-761-30590
; Sequence 30590, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FO
; TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
; FILE REFERENCE: Aecm1ca-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263,6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 30590
; LENGTH: 152
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AL022067.1

```
; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 0.95
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.1
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.3
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.2
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 7.1
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.6
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.2
; OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 1.1
; OTHER INFORMATION: NT HIT: g14759007, EVALUE 3.00e-54
; OTHER INFORMATION: EST HUMAN HIT: AUI38097.1, EVALUE 4.00e-54
; OTHER INFORMATION: SWISSPROT HIT: P10269, EVALUE 5.00e-10
US-09-864-761-30590

Query Match          48.5%; Score 97.4; DB 10; Length 152;
Best Local Similarity 90.4%; Pred. No. 1e-20;
Matches 104; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

Oy 1 CCTAGCAGCAGTAGCTCTTCAAAATCGAAGAGCTTTAGACTGCTAACCGCTGAAGAAGG 60
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 29 CCTAGCAGCAGTAGCTCTTCAAAATCGAAGAGCTTTAGACTGCTAACCGCTGAAGAAGG 88

Oy 61 GGAACCTGTTATTTTAGGGGAAGATGCTTATATGTTATATCATCATCGGA 115
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 89 GGAAGCTGTTATTTTAGGGGAAGATGTTATATGTTATTTTAGGCGGA 143

RESULT 8
US-09-864-761-7310
; Sequence 7310, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Aecm1ca-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
```

```
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 7310
; LENGTH: 520
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE: MAP TO AC010135.3
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.7
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.7
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.4
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.5
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.7
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.6
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.4
US-09-864-761-7310

Query Match          44.5%; Score 89.4; DB 10; Length 520;
Best Local Similarity 70.2%; Pred. No. 5.1e-18;
Matches 134; Conservative 0; Mismatches 56; Indels 1; Gaps 1;

Oy 2 CTAGCAGCAGTAGCTCTTCAAAATCGAAGAGCTTTAGACTGCTAACCGCTGAAGAAGG 61
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 41 CTAGCAGCAGTAGCTCTTCAAAATCGAAGAGCTTTAGACTGCTAACCGCTGAAGAAGG 100

Oy 62 GGAACCTGTTATTTTAGGGGAAGATGCTTATATGTTATATCATCATCGGAATGCTC 121
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 101 GGAACCTGTTATTTTAGGGGAAGATGCTTATATGTTATATCATCATCGGAATGCTC 160

Oy 122 ACTGGAAGTTAAAGAAATCGAGATGCAATACAGCAGAGGAGGCTTGGAAC 181
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 161 ACTGGAAGTTAAATTAATTAAGACAGCGA-AAAAAGTAGAAAAAGAGCTTATC 219

Oy 182 ACTGGAACCTG 192
    ||||| |||||
Db 220 TCTGAACCTG 230

RESULT 9
US-09-864-761-7501/c
; Sequence 7501, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL F
; FILE REFERENCE: Aecm1ca-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
```

```

; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 7501
; LENGTH: 559
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AL139245.2
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 2.5
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 3.9
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 2
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 3.3
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 3.2
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.9
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.9
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 2.2
; US-09-864-761-7501

Query Match
Best Local Similarity 42.0%; Score 84.4; DB 10; Length 559;
Matches 88; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 108 AATCCGGAATCGTCACTAGAGAAAGTAAAGAAATTCGATCGAATCAACAGTAGACGAG 167
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 559 AATCCAGAAATGTCACCGAGAAAGTAAAGAAATTCGATCGAATCAACATAGACGAG 500
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 168 AGGAGCTTCGAACACACCTGGGCGCTCCT 201
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 499 AGGAGCTTCAAAACACTGACCTGGGCGCTCCT 466
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 10
US-09-764-891-283
; Sequence 283, Application US/09764891
; Publication No. US20030077808A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC006
; CURRENT APPLICATION NUMBER: US/09/764,891
; PRIOR APPLICATION data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 10231
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 283
; LENGTH: 960
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (890)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (918)
; OTHER INFORMATION: n equals a,t,g, or c
```

```

; NAME/KEY: SITE
; LOCATION: (940)
; OTHER INFORMATION: n equals a,t,g, or c
US-09-764-891-283

Query Match
Best Local Similarity 36.5%; Score 73.4; DB 9; Length 960;
Matches 101; Conservative 1; Mismatches 47; Indels 0; Gaps 0;

QY 3 TAGCAGCAGTAGTCCTTCAAAATGGAAGAGCTTGAAGCTTACCGCTGAAGAGGCG 62
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 383 TAGCAGCCGTAGTCCTTCAAAATGCGAGACTAGACATGTTAACGCGACACAGGAG 442
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 63 GAACCTGTTATTTTAAAGGAGAGATGCTTATTAATTAATCAATCCGAATGCTCA 122
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 443 GAATTTGTTGGCCTTGATGATCAAAAATGTTCTTTGGTAAATCAATCAGAAAMGTAC 502
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 123 CTGAGAAAGTTAAAGAAATTCGAGATCGA 151
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 503 AAGACAACATCAGACAACCTCTTAATCA 531
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 11
US-10-001-835-58
; Sequence 58, Application US/10001835
; Patent No. US20020160387A1
; GENERAL INFORMATION:
; APPLICANT: Salceda, Susana
; APPLICANT: Machia, Roberto
; APPLICANT: Reclapon, Hervé
; APPLICANT: Caferkey, Robert
; APPLICANT: Sun, Yongming
; APPLICANT: Liu, Chenghua
; TITLE OF INVENTION: Compositions and Methods Relating to Ovary Specific Genes and
; FILE REFERENCE: DEX-0277
; CURRENT APPLICATION NUMBER: US/10/001,835
; PRIOR FILING DATE: 2001-11-20
; PRIOR APPLICATION NUMBER: 60/249,997
; PRIOR FILING DATE: 2000-11-20
; NUMBER OF SEQ ID NOS: 228
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 58
; LENGTH: 5434
; TYPE: DNA
; ORGANISM: Homo sapien
; US-10-001-835-58

Query Match
Best Local Similarity 36.5%; Score 73.4; DB 9; Length 5434;
Matches 104; Conservative 0; Mismatches 51; Indels 0; Gaps 0;

QY 3 TAGCAGCAGTAGTCCTTCAAAATGGAAGAGCTTGAAGCTTACCGCTGAAGAGGCG 62
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 3359 TAGCTGAGATGTCCTTCAAAATGCGAGCGGCTTGACTTACCTACCTGCTGAAAAAGAG 3418
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 63 GAACCTGTTATTTTAAAGGAGAGATGCTTATTAATTAATCAATCCGAATGCTCA 122
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 3419 GACTCTGATATATCTTAAATGAGAGAGTGTGTTTACCTAAATCAATCTGGCTGTGT 3478
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 123 CTGAGAAAGTTAAAGAAATTCGAGATCGA 157
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 3479 ATGACAACATTAATAAACTCAAGATAGAGCCCA 3513
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 12
US-09-864-761-24000/c
; Sequence 24000, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
```


TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
FILE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
FILE REFERENCE: Aecomica-X-1
CURRENT APPLICATION NUMBER: US/09/864,761
PRIORITY FILING DATE: 2001-05-23
PRIORITY APPLICATION NUMBER: US 60/180,312
PRIORITY FILING DATE: 2000-02-04
PRIORITY APPLICATION NUMBER: US 60/207,456
PRIORITY FILING DATE: 2000-05-26
PRIORITY APPLICATION NUMBER: US 09/632,366
PRIORITY FILING DATE: 2000-08-03
PRIORITY APPLICATION NUMBER: GB 24263,6
PRIORITY FILING DATE: 2000-10-04
PRIORITY APPLICATION NUMBER: US 60/236,359
PRIORITY FILING DATE: 2000-09-27
PRIORITY APPLICATION NUMBER: PCT/US01/00666
PRIORITY FILING DATE: 2001-01-30
PRIORITY APPLICATION NUMBER: PCT/US01/00667
PRIORITY FILING DATE: 2001-01-30
PRIORITY APPLICATION NUMBER: PCT/US01/00664
PRIORITY FILING DATE: 2001-01-30
PRIORITY APPLICATION NUMBER: PCT/US01/00669
PRIORITY FILING DATE: 2001-01-30
PRIORITY APPLICATION NUMBER: PCT/US01/00665
PRIORITY FILING DATE: 2001-01-30
PRIORITY APPLICATION NUMBER: PCT/US01/00668
PRIORITY FILING DATE: 2001-01-30
PRIORITY APPLICATION NUMBER: PCT/US01/00663
PRIORITY FILING DATE: 2001-01-30
PRIORITY APPLICATION NUMBER: PCT/US01/00662
PRIORITY FILING DATE: 2001-01-30
PRIORITY APPLICATION NUMBER: PCT/US01/00661
PRIORITY FILING DATE: 2001-01-30
PRIORITY APPLICATION NUMBER: PCT/US01/00670
PRIORITY FILING DATE: 2001-01-30
PRIORITY APPLICATION NUMBER: US 60/234,687
PRIORITY FILING DATE: 2000-09-21
PRIORITY APPLICATION NUMBER: US 09/608,408
PRIORITY FILING DATE: 2000-06-30
PRIORITY APPLICATION NUMBER: US 09/774,203
PRIORITY FILING DATE: 2001-01-29
NUMBER OF SEQ ID NOS: 49117
SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
SEQ ID NO 24000
LENGTH: 310
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: MAP TO AL138934.2
OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.3
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 3
OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 2.5
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 4.7
OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 2.9
OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.5
OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 2.8
OTHER INFORMATION: SWISSPROT HIT: P04027, EVALUATE 3.00e-24
OTHER INFORMATION: EST_HUMAN HIT: AW847812.1, EVALUATE 1.00e-117
US-09-864-761-24000

Query Match 35.7%; Score 71.8; DB 10; Length 310;
Best Local Similarity 66.5%; Pred. No. 1.2e-12;
Matches 103; Conservative 0; Mismatches 52; Indels 0; Gaps 0;

QY 3 TAGCAGCAGTAGTCCTTCAAAATGCAAGAGCTTAGACTTGTACCGCTGAAGAGGGG 62
DB 174 TAGCTGAGTGTCTCCCAAAACGCGAGCCTTGACTTACTGCTGAAAAAGAG 115
QY 63 GAACCTGTTATTTTATAGGAGAGAGAGCTGTATTATGTTATCAATCCGAAATCGTCA 122
DB 114 GACTCTGATATCTTAATATGAGAGAGTGTGTTTACCTAATCAATCTGGCTGTGT 55

QY 123 CTGAGAAAGTTAAGAAATTCGAGATCGAATACAA 157
DB 54 ATGACACATATAAAAACTCAAGATAGAGCCCA 20

RESULT 13
US-09-864-761-7269/c
Sequence-7269, Application US/09864761
Patent No. US20020048763A1
GENERAL INFORMATION:
APPLICANT: Penn, Sharron G.
APPLICANT: Rank, David R.
APPLICANT: Hanzel, David K.
APPLICANT: Chen, Wensheng
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL F
FILE REFERENCE: Aecomica-X-1
CURRENT APPLICATION NUMBER: US/09/864,761
PRIORITY FILING DATE: 2001-05-23
PRIORITY APPLICATION NUMBER: US 60/180,312
PRIORITY FILING DATE: 2000-02-04
PRIORITY APPLICATION NUMBER: US 60/207,456
PRIORITY FILING DATE: 2000-05-26
PRIORITY APPLICATION NUMBER: US 09/632,366
PRIORITY FILING DATE: 2000-08-03
PRIORITY APPLICATION NUMBER: GB 24263,6
PRIORITY FILING DATE: 2000-10-04
PRIORITY APPLICATION NUMBER: US 60/236,359
PRIORITY FILING DATE: 2000-09-27
PRIORITY APPLICATION NUMBER: PCT/US01/00666
PRIORITY FILING DATE: 2001-01-30
PRIORITY APPLICATION NUMBER: PCT/US01/00667
PRIORITY FILING DATE: 2001-01-30
PRIORITY APPLICATION NUMBER: PCT/US01/00664
PRIORITY FILING DATE: 2001-01-30
PRIORITY APPLICATION NUMBER: PCT/US01/00669
PRIORITY FILING DATE: 2001-01-30
PRIORITY APPLICATION NUMBER: PCT/US01/00665
PRIORITY FILING DATE: 2001-01-30
PRIORITY APPLICATION NUMBER: PCT/US01/00668
PRIORITY FILING DATE: 2001-01-30
PRIORITY APPLICATION NUMBER: PCT/US01/00663
PRIORITY FILING DATE: 2001-01-30
PRIORITY APPLICATION NUMBER: PCT/US01/00662
PRIORITY FILING DATE: 2001-01-30
PRIORITY APPLICATION NUMBER: PCT/US01/00661
PRIORITY FILING DATE: 2001-01-30
PRIORITY APPLICATION NUMBER: PCT/US01/00670
PRIORITY FILING DATE: 2001-01-30
PRIORITY APPLICATION NUMBER: US 60/234,687
PRIORITY FILING DATE: 2000-09-21
PRIORITY APPLICATION NUMBER: US 09/608,408
PRIORITY FILING DATE: 2000-06-30
PRIORITY APPLICATION NUMBER: US 09/774,203
PRIORITY FILING DATE: 2001-01-29
NUMBER OF SEQ ID NOS: 49117
SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
SEQ ID NO 7269
LENGTH: 556
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: MAP TO AL138934.2
OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.3
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 3
OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 2.5
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 4.7
OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 2.9
OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.5
OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 2.8
US-09-864-761-7269

Query Match 35.7%; Score 71.8; DB 10; Length 556;

Best Local Similarity 66.5%; Pred. No. 1.5e-12;
Matches 103; Conservative 0; Mismatches 52; Indels 0; Gaps 0;

QY 3 TGGACGACATAGTCCTTCAAAATCGAAGAGCTTTAGCTTGAACCGCTGAAGAGGG 62
DB 209 TGGCTGCACTTGTCTCCCAAAACCGCCGAGGCTTACTTACTCAATGCTGAAAAAGAG 150
QY 63 GAACCTGTTATTTTATTTAGGGAAGAATGCTTATTTATTTATCAATCCGGAATGCTCA 122
DB 149 GACTCTGCATATTTCTTAATTAAGAGAGTGTGTTTACTTAATCAATCTGGCGCTGTGT 90
QY 123 CTGAGAAAGTTAAAGAAATTCGAGATCGAATACAA 157
DB 89 ATGACAAATATAAAAACTCAAGATAGAGCCCAA 55

RESULT 14
US-09-729-674-125
; Sequence 125, Application US/09729674
; Patent No. US20010039335A1
; GENERAL INFORMATION:
; APPLICANT: Jacobs, Kenneth
; APPLICANT: McCoy, John M.
; APPLICANT: Lavallee, Edward R.
; APPLICANT: Collins-Racie, Lisa A.
; APPLICANT: Evans, Cheryl
; APPLICANT: Merberg, David
; APPLICANT: Treacy, Maurice
; APPLICANT: Agostino, Michael J.
; APPLICANT: Steininger II, Robert J.
; APPLICANT: Spaulding, Vikki
; APPLICANT: Wong, Gordon G.
; APPLICANT: Clark, Hilary
; APPLICANT: Fechtel, Kim
; APPLICANT: Genetics Institute, Inc.
; TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES ENCODING THEM
; FILE REFERENCE: 6055-64X
; CURRENT APPLICATION NUMBER: US/09/729,674
; CURRENT FILING DATE: 2000-12-04
; PRIOR APPLICATION NUMBER: 09/539,330
; PRIOR FILING DATE: 2000-03-30
; NUMBER OF SEQ ID NOS: 283
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 125
; LENGTH: 1783
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-729-674-125

Query Match 32.7%; Score 65.8; DB 10; Length 1783;
Best Local Similarity 68.4%; Pred. No. 1.7e-10;
Matches 91; Conservative 0; Mismatches 42; Indels 0; Gaps 0;

QY 1 CCTAGCAGCAGTAGTCCTTCAAAATCGAAGAGCTTTAGCTTGAACCGCTGAAGAGG 60
DB 911 CCTCGTGGAGCATCTTCAAAACGAGAGCTTCACTTCAACCATCGAGAAAGG 970
QY 61 GGAACCTGTTATTTTATTTAGGGAAGAATGCTTATTTATTTATCAATCCGGAATGCT 120
DB 971 GGGTACCTCATATTAATTAAGAGAGTGTGTTTCTGTATTAATGAATCTGGCATTTGT 1030
QY 121 CACTGAGAAAGTT 133
DB 1031 TCATATCGCAGTT 1043

RESULT 15
US-09-739-254-12
; Sequence 12, Application US/09739254
; Patent No. US20010021700A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 49 Human Secreted Proteins

; FILE REFERENCE: P2032P1
; CURRENT APPLICATION NUMBER: US/09/739,254
; EARLIER FILING DATE: 2000-12-19
; EARLIER APPLICATION NUMBER: 09/511,554
; EARLIER FILING DATE: 2000-02-23
; EARLIER APPLICATION NUMBER: PCT/US99/19330
; EARLIER FILING DATE: 1999-08-24
; EARLIER APPLICATION NUMBER: 60/097,917
; EARLIER FILING DATE: 1998-08-25
; EARLIER APPLICATION NUMBER: 60/098,634
; EARLIER FILING DATE: 1998-08-31
; NUMBER OF SEQ ID NOS: 170
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 12
; LENGTH: 2342
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-739-254-12

Query Match 32.7%; Score 65.8; DB 10; Length 2342;
Best Local Similarity 68.4%; Pred. No. 1.9e-10;
Matches 91; Conservative 0; Mismatches 42; Indels 0; Gaps 0;

QY 1 CCTAGCAGCAGTAGTCCTTCAAAATCGAAGAGCTTTAGCTTGAACCGCTGAAGAGG 60
DB 1483 CCTCGTGGAGCATCTTCAAAACGAGAGCTTCACTTCAACCATCGAGAAAGG 1542
QY 61 GGAACCTGTTATTTTATTTAGGGAAGAATGCTTATTTATTTATCAATCCGGAATGCT 120
DB 1543 GGGTACCTCATATTAATTAAGAGAGTGTGTTTCTGTATTAATGAATCTGGCATTTGT 1602
QY 121 CACTGAGAAAGTT 133
DB 1603 TCATATCGCAGTT 1615

Search completed: May 2, 2003, 18:41:58
Job time : 69.333 secs